AI in biomedicine and healthcare:

Sociological perspectives on personalized HIV therapy and skin cancer detection tools

Dissertation
zur Erlangung des Doktorgrades
der Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen

vorgelegt von
Mag. pharm. Dr. rer. nat. Renate Baumgartner

Tübingen
2024

1. Betreuer:	Prof. Marion Müller
2. Betreuer:	Prof. Jörg Strübing

Tag der mündlichen Prüfung:	27.02.2025
Dekan:	Prof. Dr. Taiga Brahm und Prof. Dr. Dominik Papies

1. Gutachter:	Prof. Marion Müller
2. Gutachter:	Prof. Jörg Strübing

Table of Contents

Acknowledgements

Some things in life I did not assume I would do once, even less so, twice. But here we are. I would like to thank the following colleagues, friends, and family for their invaluable support in the endeavor of writing my second PhD thesis: Situation Analysis Interpretation group Tübingen with Tamara Schwertl, Sara-Beth Evans-Jordan, Matthias Leger, Paulina Lutz, Maja Urbanczyk, Kiana Ghaffarizad; Schreibgruppe Tübingen with Ursula Offenberger, Birgit Kimmeler and Anne Kress; Marion Müller and her Kolloquium, in particular Hannah Bennani; ZGD with Mrunmayee Sathye, Lukas Haeberle, Fanny Weber, Aljosha Kannewurf, Alicia Protze, Darja Burljaev, Halyna Leontij, Gero Bauer, Ingrid Hotz-Davies and Regina Ammicht Quinn; Jörg Strübing and the Interpretationswerkstatt, Kevin Wiggert, the Athena Institute with Teun Zuiderent-Jerak, Lea Lösch and many more. My friends Claudia Breitenberger, Corinna Koschmieder, Anita Wieser, Timo Dufner, Daniel Lehnert, Bianca Weiss, Andreas Baur, Wulf Loh, Alisa Volkert, Philipp Rhein. My family Maria Minacapilli, Peppi Baumgartner, and Petra Baumgartner. Finally, I would like to thank myself for my persistence.

1. Introduction

1.1 About this research

The beginning of this doctoral research constituted an interest in AI in medicine and healthcare and the question of how these tools would change the context in which they were applied based on the their influence on the knowledge in the field. With the chosen perspectives of sociology and feminist science and technology studies (STS), the endeavor was narrowed down to specifically look at social categories which would be picked up, (newly) made relevant, or changed in AI-based tools. However, in 2019, it proved to be a challenge to get field access to a real-life example of AI in medicine and healthcare due to the COVID-19 pandemic. Contrary to the media hype, many AI-based tools were in research status, and access to a development team proved to be difficult. I am grateful to the colleague who suggested considering HIV-treatment optimization tools (HIV TOS) as a case study. These tools had been developed and optimized since the early 2000s and were an already established example of precision medicine (Beerenwinkel et al., 2002a). As social categories are broadly discussed within the treatment of HIV, my hypothesis was that they would also be made relevant in the AI-based tools (DAIG, 2020; RKI, 2021). However, interviews with the development team showed otherwise and brought to light the importance of not restricting the research to the AI-tools themselves, but also expanding it to the context in which they are developed and used (Baumgartner, 2023a).

This is how the idea of this study was born: To analyze which categories are made relevant within an AI-based tool in medicine in comparison with the categories made relevant in the field. Based on this assessment we can ultimately not only understand how the field shapes and constructs AI-based tools, which is a common endeavor in theory/method packages such as the social construction of technology (SCOT), but also how AI-based tools influence the epistemologies and ontologies within a field (Klein & Kleinman, 2002). For this endeavor, I also found it important to analyze a case study in which social categories are in fact made relevant for the AI-based tool itself (contrary to the former example of HIV TOS mentioned before). The outcry about racial bias that followed the market-release of skin cancer detection tools led me to believe that these tools could be a good second case study (Noor, 2020). Following a pragmatist approach, different theoretical frameworks were used to theorize my two case studies: the sociology of technology and the sociology of risk and uncertainty for the case study of HIV TOS, and feminist STS for skin cancer detection tools (Baumgartner, 2023a, 2023b; Baumgartner & Ernst, 2023). To compare the two cases in this very framework, I used the sociology of categorization, which had also informed my previous thinking about the cases. Thus, a major goal of this framework is to show how the sociology of categorization can contribute in important ways to the analysis of AI-based health technologies.

This framework is divided into two parts: The first part before the papers provides an overview of AI-based technologies in medicine and healthcare and the way they are analyzed in sociology, including an introduction to the sociology of categorization and classification. The second part of the framework (after the papers) consists of a comparative reanalysis of the case studies through this very perspective, to explore my initial question regarding the relevance of categories both in the field and for the AI-based tools.

1.2 AI-based technologies in medicine and healthcare

AI-based technologies in biomedicine and healthcare have been widely discussed since the most recent hype of AI in 2012 (Panch et al., 2019; Shortliffe, 2019; Topol, 2019a). AI in biomedicine and healthcare is being used for diagnosis, support during treatment decisions, and risk assessment or predictions of health states (Fröhlich et al., 2018; Topol, 2019a). The development and use of digital technologies have a long history within medicine (Shortliffe, 2019): Expert systems predominantly based on classical algorithms were already developed in the early 1980, and machine learning (ML) was applied in the human genome project and for precision medicine starting from the 1990s and early 2000s (Prainsack, 2017a; Shortliffe, 2019). One of these examples is the HIV TOS empirically analyzed in this dissertation (Baumgartner, 2023a, 2023b; Beerenwinkel et al., 2002a). While the analyzed case constitutes an early example of a tool with a sole focus on biological lab data, clinical decision support tools nowadays use data from different sources, including lab data, electronical patient records, etc. (Baumgartner, 2021b; Topol, 2019a). AI-based technologies in biomedicine and healthcare are described as assistants for experts or patients, and automatization is not thought to take place anytime soon (Schubert, 2022; Topol, 2019a).

1.3 Sociological research on medical AI and AI in healthcare

While for a long time in Germany, mostly technical sociologists researched AI-based technologies (see, e.g., Rammert, 1995, 1998), this changed after the AI hype of 2012. Since then, sociological literature on AI and the related processes of digitalization, datafication, infrastructures and practices around AI has been steadily growing and perspectives have diversified. Besides science and technology studies, different sociologies in Germany study AI-based technologies, including technical sociology, cultural sociology, sociology of risk and uncertainty and the sociology of comparison, valuation, categorization and quantification (Baecker, 2018; Esposito, 2014; Heimstädt et al., 2021; Heintz, 2021a; Houben et al., 2018; Nassehi, 2019; Schubert, 2022; Seyfert, 2017). Since the entire state of research would be too ample to be summarized in this framework, this section will only introduce sociological perspectives on AI used in the field of medicine and healthcare, starting with sociological research on medical AI in Germany within which I position myself.

Rammert researched medical expert systems already in the early 1980s and 1990s and wanted to find out how these systems "'socially fail[ed]'"[1] despite their "successful application" (Rammert, 1998, p. 91). Schubert and Wiggert are following the steps of this sociology of technology perspective (Schubert, 2022; Wiggert, 2021). Schubert (2022) points out the importance of studying AI empirically to show if and how an AI-based tool is productive within a particular medical field. Following a sociotechnical perspective, his team is predominantly interested in the interactions between a specific setting and a technical system, in the consequences of a new technology and in how and under which social circumstances it was developed, researching also the "interdependencies" of "knowledge, power and technology" (Schubert, 2022, p. 73). Esposito and her team are interested in ML-based prediction systems in different areas including medicine, investigating algorithmic futures applying a system theory framework. A related publication by Heimstädt et al. (2021) analyzes how contagion models circulated between the domains of public health and public safety during the COVID 19-pandemic.

Carboni et al. (2022) undertook a review of international STS and sociological literature on healthcare work digitalization. Even if their work is not specifically about AI, but about digitalization in healthcare work, it still provides a solid overview of the topics sociology is interested in when it comes to AI in healthcare, proving how

[1] German original: „trotz dieses gelungenen Einsatzes ‚sozial gescheitert'" (Rammert, 1998, p. 91).

sociological perspectives have been influenced by STS theories. Carboni et al. (2022, p. 7; Timmermans & Berg, 2003) identify the broadly used "technology-in-practice" perspective which, inspired by actor-network theory (ANT), conceptualizes technology as an actor which can construct and shape the world on its own account (Timmermans & Berg, 2003). There is an interest in researching "micropolitics of sociotechnical change" and influences on medical practices of technologies are seen "as situated within a network shaping the conditions for action and meaning-making" (Carboni et al., 2022, p. 7). Sociological perspectives see influences of these technologies as multidirectional and multifaceted (Carboni et al., 2022). Some sociological accounts see technologies as "steering", e.g., imposing institutional norms on professionals or standardizing and controlling work practices, while others see them as taking up only specific tasks within certain (managerial) logics (Carboni et al., 2022, p.8; Findlay et al., 2017; Pols, 2010). Consequences for professionals are different depending on their placement in professional hierarchies and their opportunities to co-create the technologies based on their needs (Carboni et al., 2022). Studies also address the invisible work that needs to be done in order that these technologies can be used by patients or function in interprofessional settings (Bailey, 2016; Carboni et al., 2022; Lupton & Maslen, 2017; Oudshoorn, 2008; Schwennesen, 2019). A "[r]einforcement of personal hierarchies" through these technologies and unsolved issues regarding the influence of digital health technologies on "interprofessional communication" are also discussed (Carboni et al., 2022; pp. 6, 8). Analyses see "tradeoffs of technological innovation", as some types of knowledges will be privileged (quantitative data) while others will be used less (e.g., embodied knowledge of patients) (Carboni et al., 2022, p. 5; Halford et al., 2010; Maiers, 2017).

In my work, I analyze aspects I found missing in the cited studies and used theoretical perspectives in a pragmatist way to theorize phenomena identified in the field. My own work is connected to Schubert's sociology of technology perspective aimed at understanding the function of a tool in the field, additionally using feminist-inspired approaches such as Situational Analysis to further the analysis of how power dynamics among different stakeholders provides possibilities or impossibilities of being part of technological development (Baumgartner, 2023a). Rather than focusing on the technology and surrounding practices, my analyses deal with the data, knowledge, and categories made relevant in different AI-based tools (Baumgartner, 2021b; Baumgartner & Ernst, 2023).

Medicine and healthcare are a field prone to categorization and standardization (Bowker & Star, 2000b; Epstein, 2007b). Similarly, many algorithms within AI-based tools rely on categorization (Diaz-Bone, 2018; Heintz, 2021a; Kitchin, 2014a). AI-based health technologies can influence ontologies and epistemologies in the field depending on how a problem is conceptualized within them (Baumgartner, 2021b; Kitchin, 2014b; Maiers, 2017). Moreover, medicine holds a very important role in the naturalization of social categories (Brubaker, 2009; Epstein, 2007b). Thus, the sociology of categorization can add a valuable perspective to AI research in medicine and healthcare, but it has hardly been explored so far in this field. The following section will provide an introduction into this perspective, before explaining the relevance of categorization and classification for AI-based tools, and then outline research desiderata specifically for AI in medicine and healthcare.

2. The sociology of categorization and classification

The sociology of categorization and classification builds, amongst other things, upon the sociology of knowledge and cognitive sociology (Brekhus et al., 2010; Hirschauer, 2014). Cognitive sociologist Zerubavel (1996, p. 421) famously coined the terms "lumping" and "splitting" in regard to categorization. The "active construction" of two phenomena as similar by "lumping" them into one category is similarly necessary as the construction of other phenomena as different "through splitting" (Zerubavel, 1996, p. 422). Categories can harmonize "singular and disparate phenomena" and sorting the world through classification and categorization processes reduces complexity (Heintz, 2021b).

Brubaker (2009) uses cognitive perspectives from Zerubavel (1996) and Di Maggio (1997) as the basis for his analysis of racial classification. He and fellow theorists see categories as perspectives on the world and are interested in the ways in which systems of classification and categorization work. Brubaker (2009) is also interested in the presupposed knowledge embedded in connected practices, and routines, including how they are embodied in humans, and how they help people and institutions to make sense of the world, e.g., of situations, places, and encounters. He also points out how these processes are culture-specific.

2.1 The German sociology of categorization and beyond

The sociology of categorization in Germany is influenced by cognitive sociology, sociology of knowledge and praxeology (see, e.g. Hirschauer, 2021). Hirschauer (2021) theorized categorization as a part of human differentiation processes. We distinguish, categorize, dissimilate and classify (Hirschauer, 2021; Zerubavel, 1996). Categories express the distinction between different phenomena in language, and provide "fictional groups", decreasing differences between members of the group while increasing differences to group outsiders (Hirschauer, 2021, p. 159). Categories reassure a subject of their position in relation to others based on how the subject relates to different categories of humans (Bauman, 2005). Categories also reinforce cultural order, ensuring orientation and "certainty of action"[2], and reduce ambiguity regarding how to make sense of the world (Hirschauer, 2014, p. 173; Schütz & Luckmann, 1994). This

[2] From German „Handlungssicherheit" (Hirschauer, 2014, p 173).

function and the deviations and anomalies resulting from categorization lead to the necessity of the (self-)perpetuation of categories (Hirschauer, 2014). Categories provide the following associations: categorization (allocation of objects to certain terms), identification (of actors to a subcategory), and selective social relations or associations (Hirschauer, 2014). Categorization also involves comparison: humans can be compared as humans, but are seen differently in the context of other criteria, leading to different categories of humans (Heintz, 2010). There seems to be a natural tendency to compare categories once they are formed, and to then rank them in different types of orders (Heintz, 2021b). We distinguish between similar and different facts or categories through processes of comparison. We rank them through valuation and attribute numbers to them through quantification, which gives them a different quality, e.g., of seeming more objective. (Heintz, 2021b) Social subcategories are always asymmetric, because we also position and compare ourselves to others. This can lead to strong preferences (of our "in-group") or disregard of another group. While some characteristics and behaviors are regarded as the norm, others are seen as deviant. Boundaries are constructed between the in-group and the out-group (Hirschauer, 2014; Lamont & Molnár, 2002; Wimmer, 2008). Discrepancies in categorization by oneself, by others, or by other groups can result in conflicts (Jenkins, 1997; Tajfel, 1978).

Hirschauer (2014, p. 170) sees "social affiliations" as contingent social phenomena which can be produced, used, or not used. Building on Fenstermaker and West (2001), he is interested in the "doing", "undoing" and "not doing" of social affiliations such as gender, race, and other social affiliations (Hirschauer, 2014). Hirschauer's interest lies in when and how which categories are made relevant, how subcategories are separated from each other in "boundary-making", how intense the membership of subcategories is, and which heterogenous qualities categories have (Hirschauer, 2014, p.173; Lamont & Molnár, 2002; Wimmer, 2008). A person's race, e.g., is thought to remain stable throughout life and allows a person to be part of a particular subgroup. Categories have temporalities, meaning they have varying relevance within and across time. Different "ontological registers" or frameworks can be assumed for a specific category. Race, e.g., can be seen as a bodily phenotype or as a cultural category (Hirschauer, 2014, p. 186).

Different social fields can categorize people based on different frameworks: while medicine and biology tend to think of categories as biological and are interested in biomedical difference, which according to Hirschauer (2014, p. 186) is "contingency aversive", activism against racial disparities would instead frame race as cultural, which can invite contingency (Epstein, 2007b; Heintz, 2017).[3] Different "aggregate states" which can support or diminish the stability of a category also exist (Hirschauer, 2014, p. 173). Categories can be discussed in popular discourses in the media, analyzed in scientific publications or managed through law. They are a part of cognitive schemata such as stereotypes, or practices and routines, and can be institutionalized in infrastructures and materialities such as AI (Hirschauer, 2014). A category can be solidified as an institutionalized category system such as the census, or be present in rather informal or symbolic ways (Berger & Luckmann, 1969). Different categories can be made relevant in different situations. Hirschauer (2014, p. 173), e.g., is interested in the social constructions of differences to research when and which of the different categories of a person is made relevant or irrelevant ("Welche Differenz ist wann (ir)relevant?").

Conversely to Hirschauer (2021), Bowker & Star (2000a) studied classification systems, explicit ways of categorization, with an interest in their infrastructural qualities. They define classification as "a spatial, temporal, or spatial-temporal segmentation of the world" (Bowker & Star, 2000a, p. 10). A classification for them consists of "consistent, unique classificatory principles in operation", where it is assumed that "categories are mutually exclusive" and "the system is complete" (Bowker & Star, 2000a, pp. 10-11). Even if we commonly think of classification systems in this way, the authors acknowledge that reality is messier, and classifications are ambiguous. Different classificatory schools exist, and classifications fail to reflect the world with complete accuracy. Mutual exclusivity might be hard to achieve or might not even be a goal. Thus, many classification systems have residual categories (e.g., the category "other" in the census) for everything that does not fit in the subcategories of the classification system (Bowker & Star, 2000e).

Bowker and Star (2000d, p. 322) acknowledge that classification systems can be institutionalized differently: from "small-scale seminegotiated systems" of informal classifications to "enforced universal systems such as race classification" under apartheid in South Africa. Topics within the sociology of classification research classification systems, e.g., in the political or medical domain regarding their workings, how and by

[3] This also provides a reason how categories can be stabilized in medicine through naturalization.

whom they are kept working, the work the systems do themselves, and how they change over time (Bowker & Star, 2000d, 2000b, 2000a; Lee, 1993).

2.2 Categorization and classification in digital and AI-based tools

While we have seen above that humans use categories to sort the world and act on them, categorization and classification systems are also used within digital tools, namely for saving, transferring, and analyzing digital data (Diaz-Bone, n.d.; Heintz, 2021b). Within supervised learning, e.g., the human supervisor must decide which categories are relevant to optimize the digital system based on these categories. These processes get more relevant when accounting for fairness between categories which can, e.g., involve comparing the outcome between different subcategories such as *white people* and different people of color or men, women, and other genders to check if they would be treated similarly (Hagendorff, 2019).

Bowker and Star (2000c, pp. 49, 321) already stated in the early 2000s that information systems (such as databases used for AI) “are active creators of categories in the world as well as simulators of existing categories”, and that they “are necessarily suffused with ethical and political values, modulated by local administrative procedures.” Heintz (2021b) analyzed how digitalization might influence processes of categorization, comparison, and valuation. She hypothesizes, that in digital technologies the different processes are “entangled in recursive processes with each other” and people are simultaneously quantified, categorized, compared, and put into a ranking order (Heintz, 2021b, p. 7). Heintz (2021a) also hypothesizes that the categories we know and use, such as race and gender, may lose their relevance as new categories are emerging within digital tools and will become more relevant than the categories used to date. While Heintz (2021a) based her hypothesis on recommender systems, it is then interesting to research: How do these specific workings play out in medicine which might predominantly be interested in biological aspects of categories?

2.3 Questions when researching digital categories in biomedicine and healthcare

While different sociologies and feminist STS have a long history of studying categories in biomedicine and healthcare, extensive research on the relevant topic of categories in digital health technology or medical AI is still missing (Bailey, 2017; Roberts & Rollins, 2015). Before outlining the questions to be explored in the field of medicine and healthcare from a sociology of categorization perspective, I would like to swiftly recapitulate some basic aspects of digital data to address the differences between categories both in the analogue world and within digital technologies.

The data corpus of digitalized data is limited, since not every information can be digitalized in the first place and quantifiable information can more easily be used by AI-based technologies than qualitative data (Carboni et al., 2022; Gillespie, 2014; Maiers, 2017; see also Baumgartner, 2021b). However, while recently developed technical means such as natural language processing might help to make the knowledge of qualitative data digitally available, some knowledge types like experiential knowledge or intuition might be difficult to digitalize in the first place and will not find its way into databases that can be analyzed (Carboni et al., 2022; Maiers, 2017; see also Baumgartner, 2021b).

Which data is available is relevant, because the knowledge considered for AI-based tools and the way it is processed, structured, and categorized within and through them can change how we make sense of the world. For AI in medicine and healthcare, available data might have implications on how we construct health and illness, but also human biology in general (Baumgartner, 2021b; Hofmann & Sveneaus, 2018). The way knowledge is interpreted, and categories are constructed in medicine might again influence how other fields (biology, social sciences, humanities) interpret knowledge and construct these categories (Epstein, 2007a). It can, e.g., influence which groups of humans we find it relevant to distinguish, how we construct their similarities and differences, and in which realm (natural, social) we construct these aspects (Epstein, 2007a).

From a technical point of view, some AI-based systems work based on categories which can be actively chosen (supervised learning), other systems (unsupervised learning, deep neural networks) come up with their own ways of structuring large amounts of data and it is difficult to even understand their workings (Lücking, 2020). Through digital tools, existing categories can not only be used and newly invented, but also reified, or their meaning changed (Baumgartner & Ernst, 2023; Diaz-Bone, 2018; Heintz, 2021b). In the end, this may lead to changes in epistemology and ontology in medicine and healthcare and beyond (Baumgartner, 2021b; Kitchin, 2017). The opacity of AI-based tools, being only understandable by experts, trade secret protected algorithmic systems and hardly understandable processes in technologies, all of which are termed the black box phenomena of AI, makes the question even more pressing (Gillespie, 2014; Lücking, 2020; Pasquale, 2016). This is because it is difficult to know which categories are even made relevant and only specialists such as data scientists can provide access to data sets and to the workings of AI-based systems on categories (Kitchin, 2017). However, most of the technical analysis of these phenomena do not have a social understanding of categories and how they influence our ways of making sense of the world and of acting based on them (see, e.g., Wen et al., 2022). It is thus important to research categories within and categorization through AI-based tools from a sociological perspective.

Some interesting questions to be explored here are: **Which categories are made relevant in specific tools, and how are they made relevant?** However, just researching categories in the digital tools does not provide a thorough enough perspective of the situation. Thus, I claim that a broader and more thorough analysis also needs to analyze the categories made relevant in the specific field and compare them to the ones made relevant in the AI-based tools. The following questions can then be asked: **How are illnesses and patients conceptualized in the field compared to the related AI-based tools?** Regarding the comparison between the analogous field and the AI-based tool, we also get information about which part of the problem the tool is used for, and might ask: **Which information is seen as relevant to depict the problem that the AI-based tool should solve, and how does that influence the issue at hand?** The analysis following the papers will explore these questions in the context of the case studies of AI-based skin cancer detection and HIV TOS and offer a comparison of these two cases.

3. Publications

3.1 Precision medicine and digital phenotyping: Digital medicine's way from more data to better health?[4]

Abstract

Precision medicine and digital phenotyping are two prominent data-based approaches within digital medicine. While precision medicine historically used primarily genetic data to find targeted treatment options, digital phenotyping relies on the usage of big data deriving from digital devices such as smartphones, wearables and other connected devices. This paper first focusses on the aspect of data type to explore differences and similarities between precision medicine and digital phenotyping. It outlines different ways of data collection and production and the consequences thereof. Second, it shows how these sorts of data influence dominant beliefs in the field: The field of precision medicine relying on the dominant understanding of 'genetic determinism' imported from genetics, digital phenotyping building on the logic of 'data fundamentalism'. In the end, the analysis shows how digital data informs potentials as well as challenges of precision medicine and digital phenotyping: a better health care for (some) individuals connected with individualisation and responsibilisation for all, with a prognosed shift from reactive to preventive medicine. Additionally, data-based approaches might facilitate epistemological and ontological redirections for the whole field of medicine that will also affect knowledge production and a reassessment of the value of different types of knowledge (quantifiable vs. non-quantifiable) with all its consequences. Institutionally, it might lead to shifts in distribution of power to experts in big data related technologies, i.e. private companies.

Introduction

Precision medicine (PM) and digital phenotyping (DP) are buzzwords in medicine. The promise they contain is already expressed in their naming. PM aims to offer a precise and targeted treatment for an individual or a group of people based on clinical

[4] This publication was published in British English as: R. Baumgartner. Precision Medicine and Digital Phenotyping: Digital Medicine's Way from more Data to better Health? Big Data & Society, 2021, July-December: 1-12.

data. DP promises to categorise people into clinically relevant categories for the treatment of mental diseases, with the help of digital data. Both promise revolutionary shifts in medicine and healthcare. Both approaches are data-driven, have been developed only in recent decades and hold out the prospect of changing the field of medicine on the basis of digitisable data. They are both said to transform medicine as a scientific field and as a healthcare discipline. PM claims to offer individualized treatment for specific diseases which could not only solve major healthcare challenges, but could shift the whole medical field from responsive medicine to predictive medicine (L. Hood et al., 2012). DP seeks to revolutionise mental health by collecting continuous 'real life data' from digital devices of everyday use such as smartphones and wearables. With this additional data, knowledge gaps in mental health diagnosis could be closed without burdening patients with further tasks (Huckvale et al., 2019). PM and DP, being data-based, require specific statistical and computational expertise, which results in new stakeholders entering the field of medicine.

Data today is said to be the new gold. Many publications deal with the datafication of our lives as well as with the distinct influence of data not only being collected and aggregated but also analysed by algorithmic means (Lupton, 2018b; Ruckenstein & Schüll, 2017; van Dijk, 2014). Technical breakthroughs around data saving, transfer and analysis within the last 20 years have made it possible that approaches like PM and DP have a technical basis to even be thought of and developed. Datafication and an introduction of algorithmic tools such as artificial intelligence (AI) have taken place in different fields including politics, economy, and law, but significantly also in health and the field of medicine (Topol, 2019b).

Digitalisation in medicine already started in the 1970s, then by the name of expert systems which were used to get diagnosis and treatment recommendations (Metaxiotis & Samouilidis, 2000). Over the last 15 years datafication, or 'the conversion of qualitative aspects of life into quantified data', has picked up (Ruckenstein & Schüll, 2017, p. 261). The process began with more and more patient information finding its way from paper records to electronic patient files, including the involvement of new techniques with digital output, e.g. in radiology. However, when people recently started using smartphones and wearables, engaging in digital communities, thereby producing also health-related data that can be collected, shared and analysed, datafication picked up momentum (Ruckenstein & Schüll, 2017). Thus, both technical breakthroughs and sociocultural developments such as wearable usage and sharing health-related

experiences on social media, e.g. as part of the quantified self-movement, made it possible for data-driven methods to also win ground in medicine and healthcare (Lupton, 2018b; Ruckenstein & Schüll, 2017).
All these processes have resulted in PM and DP not just remaining a vision but becoming reality and constituting two of the most spoken about data-driven approaches in medicine till date. Still, the two approaches have had historically distinct starting points. They use different types of data which is based on different ways of data production. This paper will explore what PM and DP has in common and what separates them based on the type of data they use. In the end, it will show how this data-driven aspect of PM and DP results in similar logics within both approaches, whether in dealing with mental health conditions or diseases based in the realm of the body. First, both approaches will be analysed according to the sort of data they use. In the second part, I will outline the logic behind the main arguments in the field, also based on these types of data. In the end, I will conclude with the challenges PM and DP might pose due to their role as digital data-based approaches. Comparing PM and DP might look like an odd choice, since DP is practically a part of PM, while PM is recently described as an approach to collect as much varied data as possible to compose a patient's health map. Historically, however, big parts of PM predominantly deal with genetic and genomic data, while DP is till date firmly grounded in mental health issues. Also, different stakeholders seem to concentrate on different types of data or different disease types (Mindstrong Health, 2021; Prainsack, 2017b). Contrasting PM based on genetics with DP dealing with mental health conditions, we see the similarities as well as the differences within the working logic in a field that tends to mental illnesses versus a field that predominantly deals with illnesses experts ground in the realm of the body. For this endeavour, I will focus on the following examples: breast cancer prediction, diagnosis, and treatment as an example of PM based on genetics, and the digital platform Mindstrong which claims to treat mental health issues such as depression, bipolar disorder and PTSD based on DP.

Precision medicine

PM is a medical approach that is best known for aiming to offer a specifically targeted treatment regime to a person or a certain group of people, ideally leading to a 'personalised' treatment for each person, which in the end could declare common disease labels obsolete (Ferryman & Pitcan, 2018; Prainsack, 2017b). Other terms used for this approach in medicine are 'personalised medicine' or 'stratified medicine'. The terms have appeared and been used in different historical, geographical and disciplinary contexts. I will use the term 'precision medicine' because it is the more recent term and for the sake of simplicity (Erikainen & Chan, 2019; Ferryman & Pitcan, 2018). Since 1990, PM has also been a buzzword and the closely related field of genomic science has been able to acquire considerable funding. An example is the human genome project (HGP) or the 'all of us' project in the United States (Cooper & Paneth, 2020; Ferryman & Pitcan, 2018), both with enormous expenditures. The 'all of us' project will reach total costs of about 1 billion US$ once finalised (Cooper & Paneth, 2020). PM represents an interdisciplinary endeavour combining biology, medicine, informatics, computer science, mathematics and statistics (Erikainen & Chan, 2019). Computer science, mathematics and statistics function to store and analyse data. The collaboration between human molecular genetics with modern computer science since the 1950s has been a thriving force to establish PM (Cooper & Paneth, 2020; Ferryman & Pitcan, 2018). The most prominent desiderata of PM for the field of medicine are disease understanding and prediction with a strong focus on cancer, aiming to find a suitable treatment with low incidence of adverse drug reactions (Cooper & Paneth, 2020; Prainsack, 2017b). Examples are targeted breast cancer treatment or 'personalised' antiretroviral treatment for HIV-positive individuals based on treatment optimization tools (Kumari et al., 2017; Low et al., 2018). Data about genetic predisposition, lifestyle information, clinical data, etc., all of which has to be 'structured, digital, quantified, and computable data' is combined creating a 'unique thumbprint' of a person to inform their diagnosis and treatment, also called 'personal health maps' (Prainsack, 2017b, pp. 3–4). The best-known examples till date, however, involve genetic and genomic data.

Digital phenotyping

Psychiatry got more interested in PM since struggling for a long time under the lack of objective markers for mental health conditions that science finds are a key basis for correct diagnosis and optimal treatment. Under the name of 'personalized psychiatry', and other similar labels scientists aimed to identify genomic or molecular biomarkers of mental health conditions to be able to offer targeted treatments to individuals Rüppel (2019, p. 593) describes being 'increasingly rearticulated as 'Big Data' project' (Prainsack, 2017b). DP is an advancement of this search for biomarkers that includes data collected by digital devices with sensors used in everyday life in its assessment to identify ‘digital biomarkers’ (Dagum & Montag, 2019, p. 14). DP goals are finding ‘(digital) diagnostic markers’ for mental health conditions by correlating a collection of sensor data and self-reported data and (thereby) monitoring and predicting mental health statuses based on these biomarkers, and behavioural phenotypes (Birk & Samuel, 2020, p. 1874; Dagum & Montag, 2019). The idea of DP or the digital phenotype was raised by two groups of scientists, Jain and colleagues and Torous and colleagues at the same time, both referring to the usage of large amounts of digital data to find patterns in human behaviours and traits, linking those to ‘disease phenotypes’ (Birk & Samuel, 2020, p. 1876). The assumption is that mental health conditions show themselves in digital traces of a person and ‘behaviour-expressed symptoms’ can thus be identified as ‘behavioural phenotypes’ (Dagum & Montag, 2019, p. 14; Garcia-Ceja et al., 2018). DP aims to gain data from ‘naturalistic settings in-situ, leveraging the actual real-world’ and was only possible once digital devices such as smartphones and wearables were ubiquitously available (Birk & Samuel, 2020, p. 1876). Examples of DP are the assessment of the lacking efficacy of lithium for individuals with ALS in slowing down disease progression, assessed through the analysis of online disease communities (Jain et al., 2015). Other examples are the analysis of insomnia-related tweets, and monitoring and prediction of (clinically) relevant outcomes in people with mood disorders such as major depression, bipolar disorder, and schizophrenia based on sensory measurements but also on peoples’ participation in social media (Barnett et al., 2018; Dagum, 2019; Jain et al., 2015). There is a lot of ongoing research trying to tackle different needs, e.g. the RADAR-CNS project from European Union aiming to develop an open-access platform around mobile health data, and initiatives to provide mental health treatment by smartphones to low-income populations. However, much of the research is still in an experimental stage (Melcher et al., 2020; Ranjan et al., 2019). DP is currently used mostly in the realm of mental illnesses. However, thinking of DP as part of a PM that gathers all sorts of data to compose a personal health map, it could be used for all types of diseases (Huckvale et al., 2019; Prainsack, 2017b).

As far as PM aims at incorporating all sorts of data for diagnosis and personalised treatments, DP seems to be the part of a personalisation which Rüppel describes being ‘increasingly rearticulated as ‘Big Data’ project’ (Prainsack, 2017b; Rüppel, 2019, p. 593). DP and PM are both data-driven approaches aimed at understanding diseases and offering targeted treatments, only the data they use can vary from biological, genetic and genomic data to sensor-collected data from digital devices such as smartphones and wearables (DP).

Different sorts of data − different consequences?

This section will focus on an analysis of the distinct sorts of data used in PM and DP, based on the cases of breast cancer treatment for genetics-based PM as an example of disease grounded in the biological and the platform Mindstrong for DP as an example of mental health. In the following, I will point out how the type of data influences the way the dominant beliefs within the fields of PM and DP work.

Type of data and data collection

Ferryman and Pitcan (2018, p. 3) define PM based on their ethnographic research as 'the effort to collect, integrate, and analyze multiple sources of genetic and non-genetic data, harnessing methods of big data analysis and machine learning, in order to develop insights about health and disease that are tailored to the individual.' The data involved is genetic data and a variety of other data including clinical data and lifestyle data (Prainsack, 2017b).

One method of data accumulation within PM is to start with data collection just for the means of having a large data pool. The HGP would be an example of such an approach. Human DNA data began to be collected with the promise that it would inform solutions to health issues. The project started in 1990 with the goal of mapping the entire sequence of human DNA. The main rationale behind it was to find new insights into human health and disease, which helped to secure a huge amount of funding for this project (Ferryman & Pitcan, 2018). The other approach is to conduct genome and sample analysis within pathology for specific targets. The examples on which I will focus involve the identification of specific genetic characteristics in the genome of women with a family history of breast cancer or the pathological sample of breast cancer patients, to decide upon targeted forms of therapy. People with inherited BRCA1 and BRCA2 gene mutations have an increased risk of developing breast ovarian cancers and should follow therefore more rigid preventive measures. Oncogenic human epidermal growth factor receptor (HER2) positive breast cancer patients profit from treatment with trastuzumab and lapatinib that specifically target HER2, the receptor regulating 'cell growth, proliferation and differentiation' (Low et al., 2018, p. 502).

Additionally, genomic research in breast cancer is still ongoing to identify other targets. In general, this genetic data is analysed, i.e. 'produced', at specific times by specific experts in health institutes or external labs.

Conversely, DP works with digital data collected through the usage of smartphones, computers, tablets and wearables, such as 'FitBit' or the 'Smartwatch', and other devices with digital sensors. A variety of hardware and software sensors come into play collecting data such as inertial sensors measuring walking speed, physiological sensors measuring heart rate and ambient sensors measuring temperature (all hardware sensors); but also software sensors that measure internet activity, social media presence, typing speed, etc. (Birk & Samuel, 2020; Garcia-Ceja et al., 2018). DP can be divided into behavioural phenotyping and digital biomarkers (Dagum & Montag, 2019). Behaviour phenotyping uses digital information such as location, physical activity, mood, speech patterns, typing speed and call activity, but also social media usage and search terms to search for 'behaviour-expressed symptoms' (Dagum & Montag, 2019, p. 14; Garcia-Ceja et al., 2018). It uses, e.g. passive sensing data of GPS location and call logs as a proxy for behaviourally phenotyping loneliness (Birk & Samuel, 2020). Digital biomarkers aim at measuring 'trait and state changes in neuropathology that can be indicative of disease risk, disease onset, disease progression or recovery' (Dagum & Montag, 2019, p. 14). Mindstrong, e.g. claims to have identified 'a set of digital biomarkers from human-smartphone interactions that correlate highly with select cognitive measures, mood state, and brain connectivity' that can be related to depression, anxiety, negative, positive affect and other mental conditions (Mindstrong Health, 2021). Assessments Mindstrong uses, e.g. in the ongoing AURORA study on PTSD are 'continuous-time accelerometry data, keystroke characteristics, time and duration of phone calls, time and character length of text messages, text words/ symbols used, time and number of emails, smartphone screen time, and intermittent GPS data' (McLean et al., 2020, p. 5). Saeb et al. (2016) claim mobile phone location data can be used to predict depressive symptom severity and might therefore serve as a biomarker for depression. In summary, DP is based on the digital measurement, collection, analysis and interpretation of enormously varied activities to enhance understanding and treatment of mental health conditions. Related data is collected continuously, live and in situ within 'real-life settings' (Birk & Samuel, 2020).

Data producers

Within the genetic part of PM, the processes of data production take place in highly regulated and controlled environments. Following the example of pharmaceutical labs, those facilities need to prove that they follow specific protocols even to be granted permission to produce the data. The adherence to the protocols has to be assured through different measures such as quality control, quality assurance and audits (National Human Genome Research Institute, n.d.; U.S. Government, 2021).[5] The facilities are frequently audited regarding their adherence to these protocols. Additionally, special devices used by trained expert personnel, e.g. lab technicians, are required to obtain the data. The data can only be produced by people with specialised training (lab technicians) who have access to certain facilities (labs). However, it is not just laborious but also expensive to produce this data. The necessary facilities and lab equipment are costly and so it is to follow protocols. Large teams of trained personnel are needed and have to be paid to develop and adhere to protocols and assure their quality, and are able to hold the approval status for the whole facility (U.S. Government, 2021). As a result of the resource-intensive and high-priced nature of data production, somebody has to bear the costs and the expenditures are sometimes regulated and tried to be kept at a minimum (Low et al., 2018). In the example of breast cancer, testing of the genetic properties of a patient and their sample is done once (Cedars-Sinai Blog, 2019). To summarise, the production of genetic data is initiated actively, conducted at specific time points, and is a laborious and expensive process.

For DP, the collected data is produced and shared by users of digital devices, whose live and local data can be collected continuously. The ways of collection are enormously varied, depending on the digital device and sensors used (Birk & Samuel, 2020; Garcia-Ceja et al., 2018). Often this data is collected as a by-product of practices, such as typing speed, click behaviour and not collecting personal data, as in the example of Mindstrong. Compared to the actively initiated aspect in PM, this data is provided actively (when self-reported) as well as passively (e.g. when cell phone usage is monitored) by the users. Prainsack (2017b, p. 21) calls the users of the devices 'prosumers', a combination of producers of data and consumers of information. Most of this data falls under the term big data, which has been described with properties as

---

5 For example, the mandatory CLIA certificate in the United States for laboratories performing tests of human specimen's diagnosis, prevention or treatment of disease or health problems. Genetics tests are classified as moderate or high complexity tests.

‘volume, velocity, variety, exhaustive in scope, resolution, relational and flexible’ (Kitchin, 2014b, p. 1; Rüppel, 2019). This means that it entails huge amounts of data with a high granularity. The aspect of continuous real-life data collection is named as one big advancement of DP compared to other diagnostic tools for mental health conditions (Dagum, 2019). Big data is simultaneously described as precise, because it is produced close to a person and in real life, and as messy or unclean because it might have been collected for other purposes and in an uncontrolled environment with sensors that might not be equipped to collect data in research quality (Huckvale et al., 2019; Kitchin, 2014b; Torous et al., 2016). There is a whole industry dealing (in the double sense) with health-related data called the data industry (Cosgrove et al., 2020; Prainsack, 2017b). These companies offer services around data saving, collecting, cleaning, trading and analysis (Kitchin, 2014b). The producer of the data (or ‘prosumer’) is every individual using digital technologies like smartphones and wearables, or other connected devices with sensors (Garcia-Ceja et al., 2018; Prainsack, 2017b). Additionally, any person who interacts on an online platform such as Facebook and Twitter or is simply conducting an online search provides data for this corpus of data.
The focus on the different types of data used so far in PM based on genetic data and DP reveals the following differences between PM and DP: (1) the type of data being lab-generated data versus sensor-derived data from ubiquitous available digital devices; (2) data being collected rarely versus moment by moment, live and in situ; and (3) the producers being professional labs versus users of digital devices. This next section will now show how these characteristics of data influence the logic within each of the two approaches.

Logic within the field

Drawing on Annemarie Mol’s ‘logic of care’, I will use the term ‘logic’ for a persisting rational in the field or a style that seems to be appropriate (Mol, 2008, p. 1). One could say the logic within the field is also a certain dominant rational existing in the field which is not questioned. This logic within the field seems to have different nuances based on the type of data used in the field. Historically, genetic or genomic data was used for PM, which was first expanded to clinical data in general and resulted nowadays in the usage of a whole variety of data that is helpful when composing a personal health map (Prainsack, 2017b). As I have pointed out, the most prominent examples of PM till date involve genetic or genomic data. With the establishment of genetics as a key discipline within biology, a certain dominant mindset was also established, driven by the assumption that every question regarding the functions of living organisms including human health and disease could be explained with the help of genetics. This mindset is called 'genetic determinism' (Peters, 2012). Cooper and Paneth (2020) point out that in genetically based PM, genetic determinism prevails. Weiss (2017) also calls this logic 'mendelian fundamentalism', which is an even more drastic description. Both are based on the assumption that living beings are their DNA, or differently put 'Genes R Us', and it is thus possible to find solutions against illnesses knowing about the genetics behind (Peters, 2012, p. 10). Also, if all information and logic of the DNA are revealed, it would be possible to understand how the body (and the mind) work (Peters, 2012). Similarly Weiner et al. (2017, p. 989) speak of a 'genetic imaginary' being invoked where the biological is seen as key for understanding diseases and a molecular understanding of diseases is supposed to lead to a new type of medicine. Also this logic has its roots in genetics' 'molecular vision of life' (Weiner et al., 2017, p. 999). The imaginary 'is being continuously remade and rearticulated' and gets actualised as new hopes are being constructed based on new biotechnologies such as gene editing, and the development of biological drugs (Milne, 2020, p. 103). Ultimately, the imaginary works for the persistence of the cultural power of genetics. However, in focussing on technological inventions, it is obscuring the social determinants of health and illness (Milne, 2020).

Erikainen and Chan (2019, p. 320) showed how the choice of rebranding 'personalised medicine' that happened in the U.S. policy context fell on 'precision medicine' also because the adjective 'precision' has an 'ethically neutral or even positive' connotation. I suggest that another essential argument within the field is the very aspect of 'precision', which as a characteristic not only seems to speak for the possibility of choosing a targeted treatment but surrounds genetic data like an aura. It is the conception that genetic data is precise: first because of the effort and the ways how it is measured, second because it is conceptualized as being the basis of all living beings and third having the image of being all-encompassing and everything we have to know to base our decisions on; the last two arguments being firmly rooted in genetic determinism and rearticulating the genetic imaginary (E. F. Keller, 2002; Peters, 2012; Weiner et al., 2017). It seems like these aspects give the whole field more credibility, and as if the truth is to be found in the details, and the closer we look, the closer we get to understanding it (Erikainen & Chan, 2019; E. F. Keller, 2002; Weiss, 2017). The hope to find more molecular markers for breast cancer through genomic profiling is one example of this logic (Low et al., 2018). This argument fits very well into the scientific worlds of medical science and biology, which operate under positivist assumptions. It is the mechanistic assumption that understanding a process detailed enough will reveal the truth behind it. Based on this assumption, analysing the problem more precisely will lead to better suitable treatment options simply because the premises were more precise and thus closer to the truth.[6] Hopman (2020, p. 425) in her study on forensic DNA identifies the logic of accuracy within "the search of the uniqueness of the individual". In a search for the genetic uniqueness of people, this logic results in the accumulation of data. It functions also as a 'logic of expansion' that constantly searches for new genetic '"territories" to map' (Hopman, 2020, pp. 428–429). In this constant search for more data, it permanently has to attract more money. The logic resembles very much the search for the 'unique thumbprint' to compile personal health maps Prainsack (2017b, p. 3) speaks about within PM. Also, PM is constantly aiming to attract more money, which will seem to be well invested as 'precision' implies PM will make medicine 'more effective, and thus also cheaper' (Prainsack, 2017b, p. 79). Several points of criticism can be raised to counter this logic. First, experts in genomic medicine reject that 'the more we learn about the genome, the more distant it seems to be from a role as a causative agent in most widespread diseases' and instead acknowledge the role of genomic medicine as 'a way to do science, not medicine' (Cooper & Paneth, 2020, p. 67). Medicine based on genetic knowledge might be helpful for diseases firmly grounded in genetics, however, not for other diseases. Thus, for multifactorial diseases, other measures such as public health seem to be more effective, even if they may not have the connotation of 'precision' and ultimate truth (Cooper & Paneth, 2020; Weiss, 2017). Second, critical social science perspectives would challenge the idea that a solution can only be found based on mechanistic principles. As we have seen in the genetic imaginary argument, they privilege technical solutions over taking into account systemic implications and social determinants of health. Those principles may be important for reasoning in the field. However, decisions in the field of medicine seem to be far more complex, entail much more information and

[6] Erikainen and Chan (2019) also describe how this argumentation leads to big money flows to PM. The construction of truth with the production of 'hope' within the field, in the end, is able to accumulate different forms of capital.

rely on manifold practices as studies, e.g. in medical anthropology, show (Kim et al., 2018; Spinnewijn et al., 2020).

For DP, the logic behind the argument is distinct but arrives at a similar conclusion. Generally for DP, the following goals are expressed: early disease detection and – surveillance, identifying and incentivising healthy behaviour, developing new, more targeted interventions and treatment strategies (Jain et al., 2015). Additionally, stakeholders claim that DP will be 'providing a more comprehensive and nuanced view of the experience of illness, [because] an individual's interaction with digital technologies affects the full spectrum of human disease from diagnosis, to treatment, to chronic disease management' (Jain et al., 2015, p. 462). Several experts within DP expand their hopes towards a description of how this additional knowledge gained through big data analysis will also lead to changes in classification and diagnosis and treatment of diseases 'in ways that matter most to patients' (Burtnett, 2015; Huckvale et al., 2019; Jain et al., 2015, p. 463). The Research Domain Criteria Initiative (RDoC) in the U.S. aiming at establishing a new classification system based on 'basic science' for mental health conditions is an institutional step in this direction (Rüppel, 2019, p. 571). The continuous, in situ and live monitoring of individuals' behavioural activities collected as big data are stylised as the missing piece in understanding mental health because they can be collected continuously and close to real life (Huckvale et al., 2019; Jain et al., 2015). At the same time, current diagnostic practices for mental illnesses relying on selfreporting of symptoms are framed as not conclusive enough for diagnosis (Huckvale et al., 2019).[7] One quest of DP is the search for digital biomarkers of a field frustrated by a long unsuccessful search for biological markers (Birk & Samuel, 2020; Brietzke et al., 2019). A variety of data is used, aimed at understanding mental health diseases better. This is framed as having the potential not only to guide new ways of measurement and treatment but also to change the classification of diseases and therapeutic measures (Burtnett, 2015; Huckvale et al., 2019; Jain et al., 2015). This is interesting when looking at different conceptualisations of big data. Within some fields like data science, big data is still regarded as unclean and messy and there exists a whole industry that focuses on practices of 'cleaning' and preparing big data for analysis. Stakeholders who are pro-DP mark common practices of disease identification

[7] Some practices in DP involve both DP and subjective assessments. Smartphones can be used for ecological momentary assessments which repeatedly sample behaviours, thoughts and feelings in real time (Garcia-Ceja et al., 2018).

like interviews between patients and doctors as subjective or at least not sufficient and at the same time see digital device collected data as the new Holy Grail. Thus, it seems as if lived experiences (or digital traces of everyday practices), once they are collected through digital devices which are able to provide them continuously, in situ and live, gain in worth also because it was a digital device which gathered them and because they are collected close to the individual in ‘real-life settings’ (Ammicht Quinn, 2021; Mau, 2017). Self-tracking from digital devices is framed as providing trustworthy data in contrast to the individual body’s perception which is marked as untrustworthy or at least not reliable enough to be the sole ground on which diagnosis should be based upon (Lupton, 2015).

I propose that this way of conceptualising data and guessing it would change a whole field can be termed as ‘data fundamentalism’, a term coined by Crawford (2013). Data fundamentalism is defined by ‘the notion that correlation always indicates causation, and that massive data sets and predictive analytics always reflect objective truth’ (Crawford, 2013). Additionally, digital traces are conceptualized as (digital) ‘biomarkers’ to use a common concept known in science. Accordingly, Mindstrong’s homepage indicates ‘[T]o identify the digital phenotyping features that could be clinically useful, Mindstrong used powerful machine learning methods to show that specific digital features correlate with cognitive function, clinical symptoms, and measures of brain activity in a range of clinical studies (Mindstrong Health, 2021).’ Thus, the case of DP seems to be a combination of data fundamentalism and biologisation of digital traces. It correlates digital features to states of health and illness (‘digital biomarkers’) which have big data as a prerequisite and thus, comes to the conclusion that those ‘reflect objective truth’. However, so far no digital biomarker merits the scientific definition of a biomarker, i.e. ‘a characteristic that is objectively measured and evaluated as an indicator of normal biologic processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention’ (Dagum, 2019, p. 14). Different scientists working on DP come to the conclusion that digital biomarkers are objective, and ‘reliable information on each individual’s behaviour’ (Brietzke et al., 2019, p. 223; Huckvale et al., 2019; Jain et al., 2015). Insel, e.g. compares the possibility of monitoring brain function by DP to a ‘continuous glucose monitor in the world of diabetes’ (Metz, 2018). These statements show how medical experts in the field stylise DP to provide real and reliable information about peoples’ behaviour that can be correlated to mental health conditions. They go even further when claiming they could indeed predict mental health status, e.g. relapses into depression, before individuals themselves and professionals would be able to know about it (Dagum, 2019).[8] They conceptualise big data on behaviours as a reflection of lived experiences. Actively and passively generated data from digital devices of people is framed as resulting in scientifically measured, analysable proxies for mental health and thus hold objective (and thus ultimate) truth. I would like to add that it is not just about the amount of data (which is usually important in big data) and the collection by a digital device, but data that has been collected close enough to the subject and so continuously that it seems as if it would be a direct window to peek into real life.

Looking into the depth of the type of data and in line with other authors I propose that for DP, truth is thought to be found in peoples' behaviours: what time they get up, when and how long they sleep, when and in what way they take part in social media activities, how often the text/call, how much they move, etc. From a social science perspective, it is the digital traces of people that may represent digitisable traces of pieces of everyday life practices. However, it is known that the social is digitalised only in a rudimentary form (Birk & Samuel, 2020). Also, many aspects of our everyday lives and routines are not recognised by digital devices, as our complete experiences with those practices are also not digitisable (Ammicht Quinn, 2021; Mau, 2017). There are different ways to conceptualise the collected digital features. The data about a person has been described as 'data doubles', 'data traces', 'data silhouettes' and 'digital fingerprints' (Ammicht Quinn, 2021, p. 2; Lyon, 2003, p. 22; Mindstrong Health, 2021). Loi (2019, p. 158) criticises this sort of framing that assumes being a one-to-one reflection of a person and suggests, digital information can indeed be labelled as '(digital) extended phenotype' because it may actually show health-related conditions of a person and additionally bears traces of how the user and other people retroact with it. His framing strengthens the aspect that it is also 'a collective creation, involved in social feedback loops'. Lupton claims that the 'human-data assemblages' connect body and data, but that people make their data, as well as data, makes people (Lupton, 2016, 2018b, p. 1). Still, these 'data bodies' are 'lively' in that they are 'unstable and generative' and 'lead a life of their own' (Lupton, 2018b, p. 2; Mager & Mayer, 2019, pp. 95, 98–99). These data doubles 'are not innocent or innocuous virtual fictions […].

8 The aspect of being able to predict mental health issues and the consequences thereof is an important topic that asks to be discussed, but which exceeds the focus of this paper.

They have ethics, politics' and are thus, 'no longer "doubles" […] but they are intrinsically interwoven with bodily practices and biopolitics' (Lyon, 2003, p. 27; Mager & Mayer, 2019, p. 99). Companies like Mindstrong might have less interest in tracing back data to one particular individual but more in aggregating data about a collective to gather more data for digital biomarkers. Still, the logic within the field of DP works in the way that the continual flux of 'real life data' gives us access to 'real life', so to say to a deeper truth or a truth behind it all, which is an authenticity claim. This seems to be equivalent to being closer to people. Critical data studies rather suggest that data traces and the individual behind retroact with each other and differ according to who is zooming in, how and with which epistemology and ontology (Grommé, 2016; Loi, 2019).

This assumption of having data being produced continuously and so close to the subject of interest (and thus, tied to the logic of accuracy) upholds the illusion of knowing a deeper truth about this very subject. It claims to have access to all the relevant information about the data subject needed for certain assessments. This seemingly exhaustive knowledge, continuously streamed from real life and conceptualised as objective is the basis of the assumption that we will be better in deciding which treatment options are ideal, even more so, we will be able to classify disease categories anew. In contrast, a critical social scientist perspective remarks that 'data doubles' lead their own lives and sometimes do not have that much similarity with the person from whom they are derived and therefore defy the illusion of being a window to real life (Lyon, 2003, p. 22; Mager & Mayer, 2019).[9]

Consequences and future aspects

PM will have different impacts on the field of medicine as scientific discipline and healthcare depending on what is the centre of analysis: the field's ontology and epistemology or the patient. The logic behind data analysis is driven by mathematical and statistical aspects leading to a shift to calculate outcomes for the individual from population data. In the examples of breast cancer diagnosis, the sample of individuals with breast cancer is compared to knowledge about 'genetic characteristics of a population sub-group' (Low et al., 2018; Prainsack, 2017b, p. 4). This logic of PM from the 1990s

[9] Invoking being able to monitor real life might also be due to the frustration in psychiatry of not having sufficient diagnosis and treatment tools for people with mental health conditions (Mindstrong Health, 2021).

and 2000s meant, clinically collected data of certain subpopulations served to find targeted treatment for the individual (Prainsack, 2017b). This logic stemming from epidemiology also pervades other medical disciplines. Additionally, PM has a ‘systems medicine’ approach when it comes to the object under investigation. The focus here has been redirected from molecular or cell biology to systems and models of disease pathways. This development presents an ontological and epistemological change in biomedicine (Erikainen & Chan, 2019). When centring the patient, there seems to be a twofold shift. One shift is from the individual to the population when it comes to the logic of prediction, where population data is used for. At the same time, there is a shift from population medicine to individualisation when it comes to how responsibility is framed. This means that rather than looking at population disease risk, the discipline focuses on individual risk or individual prediction. This entails a shift from the population to the individual regarding the responsibility for health. Consequently, the individual can more easily be made responsible, and population systemic aspects are not being taken into account. Erikainen and Chan (2019, p. 314) describe this with ‘responsibilization’ of the individual (rather than medical professionals) which comes from being seen as individual with autonomy first and not as part of a collective with a solidaristic arrangement (Prainsack, 2017b). Additionally, healthcare shifts from being reactive to being proactive, which can be seen by predictions that are thought to be possible based on genomic data (Erikainen & Chan, 2019; Ferryman & Pitcan, 2018). Breast cancer is an example of that when Low et al. (2018, p. 503) speak of ‘genomic profiling by clinical sequencing’ as the next step to identify cancer risks in healthy individuals. This is already a reality for people with a strong family history of cancer who are screened for BRAC1/2 gene mutations to decide about more frequent check-ups, preventive measures, or specific forms of treatment.

Similarly, for DP, experts of the field claim how DP is all about ‘person-centered care’ which also here results in more responsibility being given to each individual (Huckvale et al., 2019, p. 88). However, there are additional far-ranging consequences when analysing the processes within DP. I will now focus on those related to data types. Regarding the field’s ontology and epistemology DP even more so than PM is hoped to finally provide long longed for (digital) biomarkers for mental health conditions. The RDoC funded in 2009 by the NIMH is one prominent example for the drive for a change of the field aiming at moving psychiatry beyond so far ‘descriptive diagnosis’ for mental health conditions with a new classification system based on ‘basic science’

(Rüppel, 2019, p. 571). Big data is framed as a missing piece that might help to understand diseases in their 'expression in terms of the lived experience of individuals' (Burtnett, 2015; Huckvale et al., 2019; Jain et al., 2015, p. 463). The continuous, in situ and live monitoring of a patient's activities are stylised as the missing piece in understanding mental health.

'Empowerment of the patient' is a frequently found trope in DP, hoping that people take agency over their health which seems just another form of 'responsibilisation' found within PM in general (Erikainen & Chan, 2019; Prainsack, 2017b). However, it is still questionable who will profit from digitalised psychiatry and who will rather see the downsides. While it has been shown that access to health information, e.g. through wearable use is empowering for some people, usually those with more resources, social minorities might find it rather disempowering (Birk & Samuel, 2020; Prainsack, 2017b). DP is discussed to disregard social inequalities and conceptualise it rather as individual mental health conditions (Birk & Samuel, 2020). Real-life monitoring might bring diagnostic and therapeutic opportunities, however, it opens also the possibility to surveillance (Dagum & Montag, 2019; Lyon, 2003). Banner (2019, pp. 7–8), e.g. describes how surveillance through DP might be used 'to regulate, define and control' certain populations such as POC or disabled people and 'used to enforce neoliberal regimes of austerity', bringing more risk to historically discriminated populations than to others.

Big data in general and related data in health, in particular, is considered to be the new gold. Health data is commodified and profit is made from peoples' health data (Banner, 2019; Cosgrove et al., 2020; Prainsack, 2017b). Digital platforms are programmed with the company's purpose of collecting as much data as possible and it is a fairly frequent practice that data is sold and analysed for completely distinct purposes compared to the primary reason for data collection. Cosgrove et al. (2020) lay out that 92% of mental health apps are known to share data with third parties such as Facebook and Google without users having the choice to weigh in on that practice. This practice is called data repurposing and has been critiqued from ethical and social scientist perspectives (Kitchin, 2014b).

To recapitulate and expand: What we see in both PM and DP is the assumption that knowing the truth is enhanced through the data-driven techniques of the field. For DP, the basis is continuous, in situ and live monitoring of everyday life practices. For PM based on genetics and genomics, it is genetic data marked as holding the ultimate truth about life and being extra precise. The digital form of data is an essential aspect for the expressed logic because only then can necessary practices of collection, transfer, analysis and interpretation be possible. Digitalisation, datafication, big data and their analysis provide additional information to both fields to be able to claim being closer to their own construction of knowledge. With this newly generated knowledge, experts in the field claim that they can offer more fitting treatment options and finally, better health. Ultimately, data is an essential component for knowledge production within the field through which the field seeks to differentiate itself from other, less digital-data-prone fields. Being data-driven leads to further consequences, such as the commodification of individual's health data and how through 'personalisation' responsibility shifts to the individual.

Discussion

While both PM and DP have a focus on digitisable data, I have pointed out the differences and similarities of PM and DP related to several aspects of data use. First, I have laid out the type of data and data producers: genetic data deriving from lab facilities for PM versus big data from a variety of digital devices with sensors used in DP. Then I have shown the different systematic of data collection: active and rare data production for PM versus moment-by-moment, in situ and live data collection for DP. The focus on digital data also informs the dominant understanding within the fields of PM and DP, and results in genetic determinism and advancing the genetic imaginary for PM based on genetics and data fundamentalism in DP. Both logics share the aspect of trust in digitisable data, which is thought to hold the ultimate truth leading to better health. With this logic, these two data-intensive approaches uplift themselves over other less digital-data-prone medical disciplines. Proof for this aspect is the discursive distancing of experts of DP. They frame digital traces as 'objective' 'biomarkers' for mental illnesses against the 'subjective' standard of self-reports of mental health symptoms which can also be due to the unsuccessful search for genetic and molecular biomarkers within psychiatry (Birk & Samuel, 2020; Brietzke et al., 2019, p. 223; Huckvale et al., 2019, p. 1). In medicine, as in other scientific disciplines, objectivity ranks higher than subjectivity (Reiss & Sprenger, 2020). Information seems to gain in resilience and worth once it has been collected by a digital device (Lupton, 2015). Even more so if it has been collected close to the patient as is the case in DP. 'Real-life' data is conceptualised as being 'objective' versus self-reports being framed as 'subjective'.

Consequences might be a shift from reactive medicine to proactive and preventive medicine resulting in individualisation and responsibilisation of the patient through both PM and DP. Although DP may offer more possibilities of exchange and support for patients, the constant flow of 'real-life' data might open the door to more surveillance through connected devices, which is the other side of 24 h monitoring that, e.g. Mindstrong proposes for its users (Metz, 2018).

When data is the primary focus, the rationale is usually that more data, and sometimes more precise data as in the case of PM, is better. This is a standard rationale within big data, even though it has been criticised extensively also by the stakeholders involved (boyd & Crawford, 2012; Kitchin, 2014b). This means medicine and computer science and statistics now share this same mantra being more data is better. To satisfy this need for data within PM and DP, health institutions have to move further into the direction of digitalisation. This might bring new stakeholders, experts of data-intensive tools, into the picture who offer related products and solutions for problems. The value system of those new stakeholders will find a place in medicine and health. One result will be a further commodification of disease and health and an increased dependence on those entities that deal with data collection, analysis and interpretation, which are often private companies (Lupton, 2018b; Prainsack, 2017b; Ruckenstein & Schüll, 2017; Saukko, 2018).

Healthcare for patients may change extensively. Individualisation seems to lead to responsibilisation. Care might change from reactive to proactive and predictive. How will it affect people to know that they are continuously being monitored? To which recursive effects will this lead? Digital sociologists and STS scholars discuss that it is unclear how much patients or so-called data subjects will ultimately profit (Lupton, 2018a). Those people with more resources might indeed be empowered by tailored treatment and access to their own health data and more information. However, critical data scientists point out that social minorities might not be the ones profiting but experience the downside: being monitored can just as easily result in surveillance. People at the social margins might be depersonalised and disempowered (Prainsack, 2017b). This opens an entirely new array of questions regarding surveillance and what Foucault (2008, p. 1) called 'biopolitics', the establishment of state control over functions and processes of life. These aspects are especially salient for DP and will be even more salient once PM and DP are combined.

Being data-driven approaches, many advantages, and disadvantages of datafication and matters pertaining to big data and AI come into play. As critical scholars in social and human science have shown, only certain knowledge is quantifiable and thus digitisable. Many aspects of life cannot be translated into numbers, such as experiential, intuitive, tactile and emotional knowledge (Ammicht Quinn, 2021). If data-driven methods gain more importance in the field of health and illness, it is supposed to have huge consequences on basic epistemologies and ontologies around health and illness. Some of the non-quantifiable and nondigitasable aspects will not find their way in, e.g. decision support tools. Not only will the outcome be different if they do not inform the decision, but they will also be hidden and in the end, valued less (Ammicht Quinn, 2021; Lupton, 2015; Mennicken & Espeland, 2019). This is even more critical as both fields are rich in descriptions how data and knowledge gathered in PM and DP will change classification of diseases, identifying influences data-based approaches are expected to have on basic epistemologies and ontologies in medicine and healthcare. Therefore, critical scholars should have a close eye on these paramount changes. Prainsack (2017b, p. 188) explains how the data valuing logic in PM follows a 'tacit hierarchy of utility, with digital and computerable data on top, and unstructured, narrative, and qualitative evidence at the bottom'. A logic that was introduced by the technologies used and not by professional experts. DP in particular shows an all-too-simplified view on mental health through its search for digital biomarkers based on behaviour phenotypes. Birk and Samuel (2020) rightfully critique the usage of digital data as a proxy for social life for many reasons. First, important knowledge will not be taken into account and may ultimately be lost. Many non-quantifiable aspects have for a long time been essential for how decisions have been made within medicine. Social science shows medicine had always had a more 'human' aspect to it than science would claim. Practical knowledge, intuition and nuances in the physician–patient relationship have always been very important not just to how decisions around health and illness have been made, but also to what constitutes the professions in health care (Kim et al., 2018; Spinnewijn et al., 2020). Second, if decisions are taken based on historical data, existing categories might be reified and naturalised (Ammicht Quinn, 2021; Mau, 2017; Mennicken & Espeland, 2019). This entails also assumptions around normality and bias in algorithms which is harder to dismantle because of black box phenomena (Birk & Samuel, 2020).

Finally, all the processes described before can also be conceptualised as different aspects of biomedicalisation, the transformance of biomedicine by technoscientific interventions such as computers and genetics, with the consequences that the biological gains in importance. In the analysis before I found the following processes of biomedicalisation: a focus on health, surveillance and risk which we find in PM (genomic profiling) but above all in DP, then the 'technoscientisation of biomedicine' which can be found in both logics of PM and DP, and a change in biomedical knowledge represented by the aim for new disease classifications through data in both PM and DP (Clarke et al., 2003, p. 166). Also, the responsibilisation of the individual in both PM and DP and commodifying health data in DP fits under this umbrella term.

Conclusion and outlook

PM and DP have the potential and are already changing medicine and healthcare. The basis for these developments is also the digital data used and the logics of the field stemming from it. Since possibilities for data collection have risen, PM is seeking to include all types of data to construct a patient's 'unique thumbprint' (Prainsack, 2017b, p. 3). From the side of DP efforts exist to introduce clinical data to the phenotype analysis, so-called 'enriched data for DP' (Liang et al., 2019, p. 290). Both fields work closely together with data and computer science and share a similar big data logic. Thus, combining all data available, no matter if genetic and clinical data or 'real-life data', would be the next logical step. Some experts already depict a horror scenario of genetic determinism in psychological diseases if biological and real-life data would be collected for mental health conditions (Comfort, 2018).

Additionally, the COVID-19 pandemic has shown us how viral genetic information can be used to influence every aspect of our life, including surveillance and quarantine restrictions. Also, it is thought to change our mental healthcare, e.g. through the increased access to telehealth (Melcher et al., 2020). These transformations call for social science to engage in the field. There are many questions still to be asked: epistemological and ontological redirections in the field of medicine and healthcare, transformations in categories around health and illnesses with new technologies, changes in doctor–patient relationships and how institutions and people handle health-related data. Institutionally, it will be interesting to analyse power shifts amongst stakeholders and the changes of values and logics that accompany them. Last but not the least, addressing questions around health equality and which consequences these developments entail for population health and individual patients is crucial in these changing times.

Acknowledgements

The author would like to thank three anonymous reviewers for their very insightful comments and Tamara Schwertel, Regina Ammicht-Quinn and Ursula Offenberger for their valuable feedback. Furthermore, she would like to thank Mrunmayee Sathye and Lukas Haeberle for their help on the manuscript.

3.2 Personalized HIV treatment: Bringing marginalized patients to the forefront with situational analysis[10]

Abstract

Since the early 2000s, personalized medicine (PM) has been a much-hyped field of healthcare. HIV treatment optimization tools were one of the first successful examples of PM, and have since their development been used to find tailored and optimized treatment for HIV-positive people. In this paper on a case study of the social arena of personalized HIV therapy I show how social worlds worked on both shared and distinct goals within the arena. I highlight the simultaneous centering and marginalization to which people seeking HIV therapy were subjected discursively in the social worlds. I also demonstrate that the further patients were from practitioners' daily work, the more they were reduced to their blood samples, rather than being constructed as complex and human.

Introduction

Since work in artificial intelligence (AI) was propelled forward by major advances ca. 2012, the technology has been framed as a remedy for current challenges in healthcare (Topol, 2019b). Early examples of the involvement of AI, and more precisely machine learning (ML) in medicine can be found in personalized medicine (PM). Having been intensely hyped since the 2000s, different stakeholders were involved in PM and entered established arenas of healthcare, such as bioinformaticians and computer scientists (Erikainen & Chan, 2019). A commonly shared goal in healthcare was to ensure good health for patients. For a chronic infection such as HIV this goal is particularly ambitious. People with human immunodeficiency virus (HIV) were historically neglected by the medical community. Persistent efforts of HIV activists succeeded in influencing the research community, and the development of and access to essential medications was established (Epstein, 1996). Today, HIV treatment is framed as a success story of medicine: While an HIV diagnosis was a virtual death sentence

[10] This publication has been published as follows: R. Baumgartner. Personalized HIV Treatment: Bringing Marginalized Patients to the Forefront With Situational Analysis. Forum Qualitative Sozialforschung / Forum: Qualitative Social Research, 2023, 24(2).

through infection only decades ago, nowadays, with the right choice and availability of, and compliance with antiretroviral therapy, most HIV-positive people have a similar lifespan to those without HIV (Deutsche AIDS-Gesellschaft e.V. [German AIDS Society] [DAIG], 2020b).

Machine learning techniques were introduced to the field of HIV-1 antiretroviral therapy in the early 2000s to suggest tailored treatment options particularly for those for whom treatment selection was difficult. During the empirical research on these ML applications, I soon realized that the different stakeholders involved or standing at the sidelines of development interpreted the relevance of the tools very differently and expressed different goals while pursuing the development. Additionally, they constructed people with HIV seeking treatment in very different ways. The situational analysis (SA) framework will be shown to be well-suited for analyzing social worlds, groups who share commitment for activities, views, and/or goals which sometimes overlap with professional groups. With the help of SA, one can analyze how different social worlds interact in social arenas, where they work together on a shared problem and come to agreements about common goals. Within such arenas we also find people who are not organized in groups with shared activities, what Adele Clarke and Susan Leigh Star (2008, p. 113) called "implicated actors." Inspired by feminist philosophy of science perspective, Clarke, Friese and Washburn (2015) suggested that researchers focus particularly on marginalized actors who cannot influence how they are perceived or constructed by social worlds. Generally, the SA framework is firmly grounded in the sociology of science and technology. Both were used by researchers to analyze how technologies can change scientific research and healthcare and how this influences different parties, such as social worlds and the implicated actors involved (Clarke et al., 2003).

In this paper I present the results of an exploratory study on the productiveness of the SA framework on personalized HIV therapy in Germany as a case study in PM. I applied the social worlds/arenas concept in my analysis of how work was done within and among different social worlds within the arena of personalized HIV therapy and of which goals and activities social worlds as collective actors shared. In the first part of the paper (Section 2), I introduce the arena of personalized HIV therapy, its emergence of which was a result of work on HIV, HIV infections, and antiretroviral therapy (Beerenwinkel et al., 2002b). In Section 4 I show how people seeking antiretroviral treatment were present as implicated actors in the arena and how practitioners in the social worlds constructed them discursively in different categorization schemes which are related to the work of the different social worlds. Finally, in Section 5 the productivity of the categorization schemes for the social worlds will be discussed, as their consequences for the implicated actors, who were simultaneously centered for the shared goal of the social worlds (ensuring patient survival) and marginalized in the world's social constructions. While considering nuanced processes of marginalization of actors/actants, the suitability of SA for researching PM and digital transformations by AI will be examined from a critical stance.

HIV treatment optimization tools and the arena of personalized medicine

PM has since the early 2000s been a much-hyped field of medicine, with which scientists were able to attract considerable funding for biomedical research and healthcare. The most common goal within PM has been to use biological (mostly genetic) data to provide targeted treatment to individuals (Prainsack, 2017b). Genetic and other data about patients and illnesses have been analyzed and interpreted with the help of algorithmic systems. More recently, the applications have also been based on ML or other forms of AI which include algorithms that are trained with data to which they are exposed specifically for the task for which they are used. Due to their focus on digital data and its analysis, the development and curation of the applications new collaborations between social worlds were established. The collaborative work of healthcare experts, physicians, biologists, bioinformaticians, and others has been necessary for developing, evaluating, and refining digital tools used for PM (Erikainen & Chan, 2019). While PM has been under critique for huge expenditures not leading to an equal number of useful tools for clinical routine, HIV treatment optimization tools (HIV TOS) are an example of a successful application of PM that has been in use for almost two decades (Baumgartner, 2021b).

The ML tools which I researched are part of a family of ML-based tools based on support vector machines (Beerenwinkel et al., 2002b). They were developed and have been used within the arena of PM with the aim of treatment optimization for people tested positive for HIV[11]. The tools have been used to screen RNA data from blood samples of HIV-1-positive people for resistances to available active antiretroviral

[11] There are also ML tools for HIV-2, hepatitis B, and hepatitis C, developed by the same group of people. However, those are not part of this analysis. The researched ML tool for HIV-1 was the first of the developed applications.

ingredients. Developers of the tools also promised that the tools provide a prediction which antiretroviral drugs would be most effective against the specific HI viruses of individual patients. Due to this claim of tailored treatment the tool could be labelled as PM. The expansion of the existing social arena of HIV treatment to include the social worlds of bioinformaticians and computer scientists was a prerequisite for the development and use of these treatment optimization tools (TOS). Members of these social worlds are experts for the algorithmic reformulation of a task (e.g., to predict resistance from genetic data), and for the programming, maintenance, and servicing of the algorithmic system.

The emergence of AIDS and the HIV

In 1981 the Centers for Disease Control and Prevention noted that a few young homosexual men suffered from a type of pneumonia and a rare form of skin cancer uncommon for young, healthy people. Soon, these diseases were attributed to the early patient's "lifestyle" (homosexual sex and promiscuity) and referred to in a homonegative way as the "homosexual plague" or the "gay disease," thereby stigmatizing homosexual and bisexual men and men having sex with men. At first there was no treatment available, and politicians of many countries (including Reagan in the US, Thatcher in the UK, and Gauweiler in Bavaria/Germany) were not interested in providing a treatment but rather in further marginalizing affected groups with high incidents of HIV/Acquired immune deficiency syndrome (AIDS), such as homosexual men, sex workers, and drug addicts. This resulted in HIV-positive individuals and whole communities (lesbian, gay, bisexual, and transgender [LGBT]) rising up in activism for research on and a suitable treatment for HIV/AIDS, and resisting their marginalization (Beißwenger & Höpfner, 1993). In 1982, the disease got its official name, AIDS, because it weakens or destroys the immune system and makes the body vulnerable for various types of infections. Shortly afterwards, as AIDS cases were reported in many countries and on all continents, AIDS was given the status of a world health problem, which it remains to this day (Epstein, 1996).

In 1983, the virus that causes AIDS was discovered and later named Human Immunodeficiency Virus. As a retrovirus, HIV's genetic information is contained on a single strand of RNA. The virus enters the cell, forcing it to transcribe the viral RNA into the cellular DNA. The cell then produces new viruses according to the rival HIV-RNA specifications. HIV primarily affects what are called CD4 lymphocytes/helper cells, which are essential for a functioning immune system. Every day, millions of new viruses are produced in the organism of an HIV-infected person, and just as many viruses are destroyed by the immune system through the destruction of helper lymphocytes. This can be compensated by the immune system for many years, but eventually the decrease in the helper lymphocytes weakens the body's own defense system. HIV reduces the body's ability to fight infections and disease and can then lead to death via opportunistic infections and tumors (Gallo & Montagnier, 2003; Schmid, 2018). In the late 1980s it was found that HIV is especially prone to mutate. This remains one of the major challenges in treating HIV infections (Epstein, 1996).

HIV/AIDS-treatment

The first medication for HIV-infected people, called azidothymidine (AZT), only became available in 1987 and was at first used as a single drug treatment. Besides the severe side effects, the problem of this single drug treatment was that mutation-prone HI viruses developed drug resistant mutations. Once the virus is resistant to the drug, the treatment stops working and the immune system deteriorates, leading to the infections and tumors mentioned above (Epstein, 1996). Over the following years, several other drugs similar to AZT were released. Initially, these were also administered as single drug treatments sequentially in response to the development of medication resistances. Soon, people treated in this way ran out of medication options. In the 1990s, a treatment approach was trialed where three or more drugs were given in combination, called highly active antiretroviral therapy (HAART). Only with the advent of HAART in the mid-1990s did long-term effectiveness of HIV/AIDS therapy become possible by pushing the viral load to a very low level. With this new form of therapy resistance-inducing mutations of the virus against available antiretroviral drugs could be stopped. HAART has prolonged the lives of many HIV-infected people and its widespread use has resulted in HIV being rather a chronic disease than a constant threat to life. Thus, the main goal of antiretroviral therapy to date is that viral load is decreased as much as possible, ideally to an undetectable level, which also diminishes the chances of HIV transmissions. Even then, HIV resides permanently in the body. Hence, antiretroviral therapy is to be taken lifelong and without interruption. Both therapy interruptions and underdosing of the medication carry a high risk of treatment failure due to the development of resistance (DAIG, 2020b). Therefore, a patient's compliance in regular and lifelong drug administration is required for a successful antiretroviral therapy. In the mid-2000s, single-tablet regimens were approved, in which HAART, consisting of several active ingredients, was combined in one daily tablet (European Medicines Agency, 2018). This represented a breakthrough in HIV treatment for the physicians and patients involved, making it significantly easier to follow an active treatment regimen. Patients' compliance was better achieved with the simplified treatment regimen. This led to increased treatment success. Nevertheless, virologist Bio1 in my study reported there were around 10% HIV-positive patients for whom resistances complicate the choice of treatment (Bio1, r.282ff.[12]). Here, different experts from different social worlds come into play, as we will see in the following analysis.

Social worlds, social arenas, implicated actors, and discursive constructions

The key focus of SA is "interpreting the situation per se" (Clarke et al., 2018, p. 27). Theories of William I. Thomas, George H. Mead, John Dewey and especially Herbert Blumer led to the concept that "a situation is a gestalt greater than the sum of its parts" (Clarke et al., 2018, p. 71). "This *invisible agency of the situation per se*" is "the *momentum of the relationality among the different elements of the situation*" (p. 70) and is the key interest to be explored during analysis. The concept of social worlds has been used "since the early days of Chicago-style interactionism" (A. L. Strauss, 1978, p. 119) and is "the conceptual infrastructure of situational analysis" (Clarke & Star, 2008, p. 114). As "social wholes" (p. 115) from the Chicago School of Sociology in the 1950s and 1960s, they were used to study work and professions focusing on shared discourses, instead of geographic boundaries, also considering the interaction and discourses between them. Within the modern version of social worlds theory, social worlds "generate shared perspectives" (Clarke & Star, 2008,) as a basis to work on shared activities and to "achieve their goals" (Clarke & Star, 2008) with shared resources of various kinds (see also Clarke et al., 2018, p. 71; A. L. Strauss, 1978). Social worlds have at least one primary activity, specific sites, and involve technology to pursue their activity (Clarke et al., 2018; A. L. Strauss, 1978). Social worlds are procedural, i.e., they do not exist by themselves, but only become visible in practical consequences: They come into being and exist only through the joint commitment of their

[12] With the numbers the lines in the transcript are indicated, and "ff." is used for following lines. I translated all transcripts from German to English.

members regarding their core activity and the mutual reference to each other. It is through engaging and participating in social worlds and arenas that individual and collective identity is formed (Clarke & Star, 2008). Especially for modern societies, it is constitutive that people belong to many different social worlds, i.e., have multiple memberships, which are susceptible to change. This creates overlaps at the edges of and differences between social worlds. Thus, social worlds intersect, segment into subworlds and are fluid (Clarke et al., 2018; A. L. Strauss, 1978, 1984). "*Social worlds* are groups of varying sizes that generate life of their own" such as a discipline or a profession (Clarke et al., 2018, p. 71). In this perspective, "society as a whole can be conceptualized as consisting of a shifting mosaic of social worlds that both touch and interpenetrate" (Clarke, 1998, p. 16). Important characteristics of social worlds are Mead's understandings of a shared perspective and commitment to a situation/arena. When social worlds grow and intersect into arenas, "their joint courses of commitment and (inter)action are articulated through discourses" (Clarke & Star, 2008, p. 116). Following Mead, one can conceptualize them as "*universes of discourse*." These discourses are "in collective, material action" (p. 115).

Social arenas are "composed of multiple worlds organized ecologically around issues of mutual concern and commitment to action" (p.113). From an analytical perspective, with the concept of social arenas one can analytically determine the processes of exchange and the relationships between social worlds and subworlds. It is a scalable concept, with which larger and smaller arenas can be investigated in and between social worlds (A. L. Strauss, 1993). Being part of Chicago School interactionism, social worlds/arenas theory also lends itself to the analysis of differences, e.g., of perspectives between social worlds. Anselm Strauss saw social worlds and arenas as "significant sites of *negotiations*" and "negotiating" as "major social process" (Clarke et al., 2018, p. 73). Clarke et al. described arenas also as "*discursive* sites" that can last over time and will then be "characterized by multiple, complex, and layered discourses" (p. 73). As "sites of contestation and controversy," arenas are a particularly helpful tool for analyzing "heterogenous perspectives, and position on key elements, and to see power in action" (p. 73). Within Clarke et al.'s framework, "work activities, organization, and discourses" (p. 75). i.e., practices of and action within social worlds rather than individuals, are the focus of analysis. Further, both human and nonhuman elements of a broader situation can be considered and have agency. Working with different analytical maps is an essential feature of SA. *Social worlds/arenas maps* can be used to "lay out the collective actors, key nonhuman elements, and the arena(s) of commitment and discourse within which they are engaged in ongoing negotiations—[i.e.,] mesolevel interpretations of the situation" (Clarke & Star, 2008, p. 128). Since the 1980s, social worlds theory has also been used in STS and studies of social worlds in the life sciences, e.g., to explore nonhuman actors, tools and infrastructures (Clarke & Star, 2008).

From various studies with social worlds theory (many focused on scientific work practice), what Blumer (1969) called "sensitizing concepts" (p. 147) have been generated. According to him, these concepts are not clear definitions and specifications (definitive concepts), but rather suggest a direction for the investigation. Clarke and Star (2008, p. 117) named a whole toolbox of sensitizing concepts that can be used for analysis in social worlds theory "to think about the relational ecologies of social worlds, arenas, and their discourses." One sensitizing concept which has a prominent place in situational analysis is that of *implicated actors/actants.* This concept was a result of Clarke's empirical research (Clarke & Montini, 1993) and can be a useful tool to analyze relative power in social worlds/ arenas: "It focuses on the situatedness of less powerful actors in a situation and the consequences of other's action for them, raising important issues of discursive constructions of actors and nonhuman actants and their consequences for those actors" (Clarke et al., 2018, p. 76). Implicated actors are constructed by others for their own purposes. Clarke and Star (2008) described "at least two kinds of implicated actors" (p. 119) which are relevant for the study reported here: "those who are physically present but are generally silenced/ignored/made invisible by those in power in the social world or arena," and those who are "*not* physically present in a given social world but solely discursively constructed and discursively present; they are conceived, presented, and perhaps targeted by the work of the arena participants" (Clarke & Star, 2008). Implicated actors play no active part in the creation of their representation within these social worlds or arenas, their perspective is not seen as relevant by the members of the social worlds, and they are not asked to participate in the arena (Clarke & Star, 2008). Clarke and Star formulated the following analytical questions for implicated nonhuman actors: "who is discursively constructing what, how, and why?" (2008), which I will show are also productive when studying implicated human actors. In the current analysis, the concept is used as a starting point to develop a suitable version of *implicated human actors* for this particular case study.

For a more detailed analysis within the social worlds, the concept of *discursive constructions* was used. This term was adopted by Angelika Poferl (2004), Rainer Keller, Andreas Hirseland, Werner Schneider and Willy Viehöfer (2005) and is connected to Peter L. Berger and Thomas Luckmann's sociology of knowledge (1969) and Michel Foucault's discourse (1974) (see also Bosančić & Keller, 2019). Similar to grounded theory methodology (GTM, A. L. Strauss & Corbin, 1994) and SA, it is connected to pragmatism insofar as problems in societal action and thinking are considered starting points for discursivation of phenomena and materialities. Saša Bosančić and Keller (2019) spoke about discourses as connected to the materialities of existences. They "produce hard facts" and "bring order" to the field (p. 2). The focus on analysis of discursive constructions in the methods mentioned above resonates with the power-critical orientation Clarke built into SA, where the discursive work which happens in social worlds is analyzed with a view toward revealing power asymmetries. [12]

The following section contains my analysis of the project of the ML tools mentioned in the beginning of the chapter, which I examined as an example of an arena of PM for HIV treatment in Germany. I aim to illustrate how SA can be used in the course of a research project to analyze social worlds and implicated actors within an arena, and how the concepts sensitize the view of the researcher. As Clarke et al. noted, "SA itself builds on the genitivity of the situation through analysis" (2018, p. 70). SA, rather than the more traditional social worlds/arenas analysis of the arena of PM for HIV treatment, was chosen because of two distinct characteristics. Firstly, with SA the researcher can focus their analysis also on nonhuman elements. Even if those are not central in this particular analysis, keeping nonhuman elements, such as the HI viruses, blood samples, or the ML tools in mind during analysis was very important. Secondly, through Clarke et al.'s approach, one can consider that even if social worlds are sometimes represented by a few voices of their members, those still speak for the social world as a whole. Hence, the analysis is not one of individuals, but one of collective "work activities" and discursive constructions that members of these social worlds develop.

I will also explore how doing different iterations of a social worlds/arenas map can sharpen the analyst's sensitivity to different aspects within the analysis. I will illustrate this by presenting my analysis of (shared) goals and activities, and categorization schemes that the social worlds of virologists, bioinformaticians and physicians are discursively constructing, while searching for a suitable treatment against HIV. The analysis is based on semi structured expert interviews according to Michael Meuser and Ulrike Nagel (2010) with one physician (Med1), two virologists (Bio1 and Bio2), and two bioinformaticians (DS6 and DS7). They were chosen as experts in their field (all) and because they had published about the researched ML-tool (all virologists and bioinformaticians). Apart from the physician, all the interviewees were closely involved in the development of these tools. The interviews were transcribed verbatim before being coded based on GTM. Several documents on HIV treatment, homepages of the researched tools and the related social worlds, publications by them and guidelines for HIV treatment were coded and used in the SA. Social worlds and their organizations, activities, technologies, and goals were identified following the guiding questions that are sketched out under Section 4.1. Messy situational maps, relational maps, social worlds/arenas maps and mind maps detailing processes within the arena were used in the analytical process. Social worlds/arenas maps were adapted for the work in this case study.

Analysis

Social worlds in the arena of the researched ML-based HIV TOS development

When identifying social worlds, one usually focuses on analyzing the following questions: *Which primary activities take place? Where do they take place? Which technologies are used? Which organizations are involved?* An overview of these aspects is shown in Figures 1 and 2 and the analysis below. The initiative for the development of the researched ML TOS for HIV came from the clinicians who worked with HIV patients (Bio2, r.62ff.). However, virologists and bioinformaticians were central actors in the development of these tools and worked very closely together, with clinicians as advisers (Bio2, r.80ff.).[13]

[13] The social worlds of computer scientists, although involved in the social arena of the development of these HIV TOS, e.g., in taking care of day-to-day system administration, did not play a central role in this case study and will not be present in the following analysis.

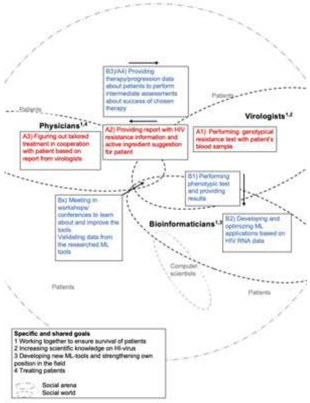

Figure 1: The goal and activities map of the researched ML-based HIV TOS. Please click here for an enlarged version of Figure 1.

The primary activity of virologists Bio1 and Bio2 took place in laboratories at the University Clinic of Düsseldorf[14] and consisted of sequencing the HI-viral RNA derived from the patient's blood sample to analyze the viral RNA for mutations and determine the presence of resistances. According to a literature search, two different sorts of tests can be used. The most common test was and still is the genotypical resistance test, which compares the HI-viral RNA with known other HIV, searching for mutations and predicting susceptibility of HIV to antiretroviral medication based on "interpretation systems"[15] (Deutsche AIDS-Hilfe e.V. [DAH], 2010, n.P.; Gemeinsamer Bundesausschuss [G-BA], 2004, p. 30). The second option is a phenotypical resistance test where the HI-viral RNA is checked in an in-vitro titration for existing resistances against known active antiretroviral ingredients. Different tests were carried out in different situations and different information was made available to different social worlds. Bio1 stated that virologists normally provide a report to the physician outlining the active antiretroviral ingredients to which the specific HIV is susceptible. The scientific basis of the report usually consists of results of genotypic resistance tests, which is the more economic option (DAH, 2010). Physicians used this to determine the most suitable antiretroviral therapy (Bio1, r.302ff.). Virologist Bio1 (r.78ff.) spoke about the advent of ML-based HIV TOS development. It started when they performed phenotypical resistance tests on the blood of therapy-failure patients. The resulting information—which HI virus was resistant to which ingredient—was the key training data for some of the earlier ML tools. Here the bioinformaticians, as experts in developing

[14] City names were changed for anonymization.

[15] All translations from German are mine.

algorithmic applications for HIV TOS, came into play. Their primary activity, taking place mainly at the Munich Institute for Computer Science, was the development and optimization of ML applications based on the data the virologists obtained in the cell culture in the lab. This data was used as training data for the ML algorithms in the tools. Once the bioinformaticians and members of other involved social worlds were satisfied with the performance of the ML system, the application was released and freely available for everyone to enter HI-viral RNA sequences and obtain predictions of resistances regarding the specific RNA entered (DS6, r.1370ff.). While developing the researched ML tools, bioinformaticians had to coordinate their work closely with the virologists who are the experts on the virus, e.g., they needed to negotiate common definitions what "therapy-failure" means and how that could be translated to the algorithmic system (Bio1, r.446ff.). The bioinformatician recognized that to increase acceptance and the use of the tools, it was crucial for users (virologists and medical professionals) to understand how the tools worked, how the results are obtained, and how to use them properly[16]. The interviewed virologists, being expert users, usually evaluated the suggestion of the tools with their own scientific experience or scientific literature on the subject. The bioinformaticians also recognized that the tools must be adapted to the expectations and needs of the users (the community), before collaboration partners could be won (DS6, r.720ff.). Yearly workshops and conferences were organized for an exchange between different social worlds. In these arenas, physicians, virologists, and bioinformaticians gathered in person to discuss the tools more thoroughly, get feedback for further optimizations and developments, and exchange knowledge about functioning and correct use of the tools (DS6, r.456ff.). The attendees used these workshops to scientifically validate the outcome of the researched tools by discussing specific patient cases.

The shared goal that connected the social worlds can best be described with the words of an interviewed virologist: "Ultimately, we want to ensure the survival of the patient [...] There should be a quality of life, right? And these- these are challenges where we have to have the courage to consider things that are not yet mainstream" (Bio2, r.366ff.). This means that actors in the medical field, biology, and bioinformatics worked together for the medical goal of ensuring survival of patients (Figure 2). One

[16] Here, the bioinformaticians acted in line with current ethical discourses in the development of medical ML tools, which are also concerned that the outcome produced by these tools should be explainable (Baumgartner, 2021a).

way this was achieved was by strategically linking existing knowledge from different research areas, and recruiting new stakeholders—in this case bioinformaticians—in order to be able to develop the ML tools. Further social-world-specific goals were to gain knowledge about virological diseases (virologists), to develop new algorithmic applications for use in other medical areas, e.g., cancer research and infectious diseases. This latter goal was mostly interesting for bioinformaticians, who aimed to apply their algorithmic systems in various medical areas (DS6, r.98ff.). In my analysis, I also identified the personal and/or professional goals of building a scientific reputation (all researchers).

Despite the impression of concerted cooperation in the pursuance of common goals described above, reception of the researched tool was not universally enthusiastic. Its relevance was evaluated very differently by members of various social worlds. While the virologist and the bioinformatician saw a relevance in the tools they invented, the physician I interviewed (Med1), who was not part of the development, was not acquainted with the tools.[17] Med1 understood their own role as consisting in working with the patient and finding a suitable therapy together, and acknowledged that expert knowledge regarding HIV would come from the virologists (the main users of HIV TOS). Med1 (r.1311ff.) did not identify any problems in finding a suitable and active therapy, even less so since single-tablet regimens became available. In summary, I have used the SA framework to demonstrate how applications based on AI and ML, such as the researched tool in the field of PM, were constructed within and through arenas composed of heterogeneous actors committed to action on a core issue. I have also shown that while commitment to common goals is characteristic of the work done in social worlds, individual actors experienced different levels of commitment.

Marginalized while centered actors ("patients")

People with HIV who seek treatment were the connecting elements between the interviews, which makes sense considering the shared goal of the arena was to ensure their survival. The actors in the social worlds centered this group of people through this goal

[17] It is important to distinguish between the physician I interviewed who did not use HIV TOS and did not know a lot about them and the clinicians that have indeed been part of the development of the tools mentioned by the interviewed virologists and bioinformaticians. This second group will be interviewed in the future. Accordingly, in the social worlds/arenas map on Figure 1 a small overlap with the social world of bioinformaticians (for the few clinicians working on the development of HIV TOS) and a bigger one with virologists is depicted.

as pivotal actors within the arena. The existence of the patients has been a prerequisite for the existence of the arena. As implicated actors, the "patients"[18] were present in diverse ways (Figure 2) and discursively constructed by actors in all examined social worlds (physician, virologists, bioinformaticians). However, the ways in which they were discursively constructed shows they were objectified and marginalized, which manifested as nuanced processes across social worlds. Two examples of this marginalization were being excluded from shaping the discourses about their own group and from being part of a decision that is above the level of their individual treatment.[19] Thus, the patients found themselves in a position of less power compared to the actors in the social worlds within the arena.

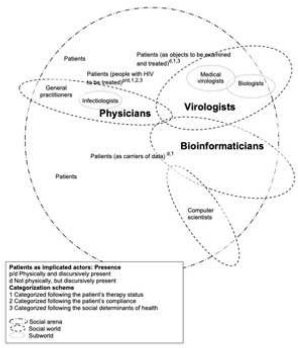

Figure 2: The arena of personalized HIV therapy. Please click here for an enlarged version of Figure 2.

It was in the interview with the physician (Med1), who had the most real-life contact with people seeking HIV assessment or treatment, that the patients seemed most present as individuals. Physician Med1 specifically referred to compliance, stigmatization, pregnancy, and the patient's cultural background. For the physician, who saw people with HIV during regular control examinations, the patients were present physically and discursively, while technical applications played almost no role. Conversely,

[18] I use the word "patients" as an in vivo code, when the social worlds are cited and when referring to people who started HIV treatment and were thus seen by different social worlds as a patient. To highlight a critical stance towards the term, the fixed role it assigned to people and the implications this role has had, I replace it with "people with HIV that seek treatment" or similar wording in other cases.

[19] People with HIV coming for treatment were not interviewed in this study. From my online and literature searches I assume, that some of them were organized in self-help groups. However, these groups were not mentioned by my interviewees and were not part of the development of the studied HIV TOS.

bioinformaticians had no physical contact with people with HIV but saw them in terms of their data. Bioinformatician DS6 discursively constructed them as carriers of data:

> "So that means, when you now have a patient, normally you have one sequence per patient, right? With Sanger[20] you have this one particular consensus sequence that you see. And with NGS you have several sequences per patient. Because you look deeper and see different virus variants" (DS6, r.632ff.).

Seeing patients primarily in terms of their viral sequences, the professional group of bioinformaticians was the furthest away from people with HIV among the social worlds, followed by the virologists (Bio1 and Bio2). With the interviewed virologists, a more differentiated picture emerged. While Bio1 talked about people who had come for HIV assessment or treatment mainly in terms of the objects to be examined (i.e., blood samples which were subjected to resistance measuring), Bio2 presented a more holistic view on the disease, using references to popular culture and well-known personalities. When referring to clinical patients or those infected with HIV, they especially discussed people in terms of their cultural background, gender, e.g., as women and as pregnant women, and acknowledged their stigmatization, picking up similar topics as the physician.

Categorization schemes of patients in different social worlds

In all examined social worlds, I could distinguish between different categorization[21] schemes used to cluster patients into types. Members of each social world, depending on the internal logic that had been developed/negotiated, used certain categorization schemes fully and others just partially or not at all. The first categorization scheme related to the therapy status, the second one to the patient's compliance. A third scheme was based on social determinants of health that were made relevant for HIV-positive people (pregnancy, cultural background).

---

[20] Sanger is the traditional way of sequencing. NGS is the newer method, enabling a deeper look into the virus and a more complex picture of resistance probabilities. An ML tool developed by the same working group is available for NGS data (Bio1 and DS6).

[21] The boundaries of a category, as opposed to a classification, may be "fuzzy," and "membership [is] based on generalized knowledge and/or immediate context" (Jacob, 2004, p. 528).

Categorization based on patient's therapy status

Categorization based on the patient's therapy status was strongly prevalent in the accounts of the virologists, who regarded people with HIV as objects of treatment. Resistance tests were carried out at the beginning of treatment, when the viral load increased, and in the context of changing therapy. These tests provided essential information for the choice of the optimal antiretroviral therapy.

> "That means we have these three cases: We have the therapy-naive, we have the failures whose viral load goes up, and we have those who do not have a viral load but want to change the therapy, for whatever reason. Those are the three cases" (Bio1, r.445ff.).

Therapy-naive-patients: If a person's HIV test was positive, a resistance test was carried out in a virological lab using the person's blood sample. At this point the whole range of medication was still available for treatment in most cases; only about 10% of the therapy-naive-patients were infected with HI viruses that were already resistant to some drugs (so-called transferred resistances). The virologists examined the HIV in the blood and informed the physician about which active substances the virus is not resistant to. A therapy suggestion was made by the physician to the patient. The physician was the only one of the three expert groups who knew about a possible co-medication and other characteristics of the patient to be considered. The ultimate therapy decision was taken by the physician together with the patient.
Therapy-failure-patients[22] *or the problematic/difficult case:* According to my interviewees and in accordance with the Austrian-German guideline for antiretroviral therapy of HIV-1 infection (DAIG, 2020b), the viral load and the number of CD4 cells in the blood of HIV-positive people was measured every three to six months (the frequency depending on compliance, as assessed by the physician). As soon as the viral load increased, a resistance test was carried out. One reason for therapy failure could be virus mutations leading to resistance to the current drug therapy. Patient non-compliance was considered the main reason for this, i.e., "because the drug wasn't taken properly. That is the classic case. If the dosage isn't right, if it's underdosed, then the virus has a chance to replicate, to multiply, and that's when the mutations come about" (Bio1, r.202ff.). The virologists referred to these as "therapy-failure-patients," while the physician said, "the patient had a virological failure" or called them "problematic

[22] In German Bio1 said "Versager," which I translated to "failure" or "loser."

cases," and one bioinformatician spoke about "difficult cases" (Bio1, r.79; Med1, r.1104, r.419; DS6 r.1197). Asked for other possible explanations for the development of resistance, bioinformatician DS6 (r.1093ff.) replied: "That depends on many factors, so it can be that the medication is not taken regularly, or that the other medications had side effects, or that they had a gastrointestinal illness and therefore did not absorb the medication well." The physician and the virologists pointed out that it is usually older patients who had been in antiretroviral treatment for many years who experience therapy failure. A lack of effective treatment regimens over the course of decades meant many active ingredients had been tried out (also as single treatments), allowing a person's virus to develop resistances against many of the ingredients. This meant that ultimately, few active ingredients remained available to them, and it was difficult to offer a functioning HAART. However, at "maybe 1%" (Med1, r.400ff.), those cases were extremely rare.

Change-of-therapy-patients: A third category consisted of patients whose therapy was being changed without having experienced a therapy failure. This was happening either at their own request due to co-medication, side effects, or drug interactions, or because simpler therapy regimens, e.g., single-tablet therapies, were available. In this case, there was no increase in the viral load, but a preventive check was carried out to determine whether there was any resistance to the drug under consideration.

This categorization scheme was based on the logics of the social worlds[23] that were invested in treating the chronic disease of HIV-positive people and ensuring their survival. Categorization into the patient types introduced above was most relevant for the work of virologists, and the full scheme was constructed by them and the physician (though one of the virologists and the physician did this descriptively and not in the exact terms I have used here). By categorizing patients into types, different practices and actions were triggered. The scheme also illustrates how the virologists saw patients as subjects in different treatment status groups. Since its approval in 2007, and with an increasing choice of different marketed variants within the last 5 years (Med1, r.192ff.), the single tablet remained the regimen of choice for most people treated for HIV infections in Germany. With this regimen good outcomes, i.e., low viral loads, could be obtained. This left only approximately 10% of patients that represented

[23] "Logic" is here used in the sense of a persisting rational or an appropriate style in the field as published by Mol (2008, p. 1).

somewhat of a challenge for this social arena of PM, and 1% who represented truly difficult cases. This patient type seemed to be the most labor intensive one, mentioned by Bio1 as posing a particular challenge to their expertise. A "therapy-failure" was the type where members of all social worlds came into play, and it tied them together more strongly, uniting them in a need to rise to the challenge and collaborate to find a suitable therapy. This was also the only type of patient that was constructed discursively by all three social worlds, despite with slightly different meanings. While the virologists spoke of "failures," the physician and the bioinformatician spoke less judgmentally about "problematic" or "difficult cases," respectively (Bio1, r.445; Med1, r.1104, r.419; DS6 r.1197). These 1% of patients who had had a long history of therapy and who had already developed many resistances were the most demanding cases not only for virologists but also for bioinformaticians. The bioinformatician speculated that especially in such cases, the ML application could deliver better (or at least equally good) results compared to older rule-based systems (DS06, r.1202ff.). Ensuring the survival of this type of patient seemed therefore to be the main jointly constructed reason for the (ongoing) development of the researched family of ML tools.

Categorization based on patient's compliance

The second categorization scheme reflected patients' compliance with their treatment regimen. It followed mostly the logic of the physician but was discussed by all three social worlds. Patient compliance was considered to be the decisive factor for therapeutic success. Since HIV therapy is life-long, durable patient compliance was and still is crucial. The physician constructed patients in this categorization scheme in terms of compliance-types:

> "For some this is totally unproblematic, and it helps because they say: Yes, well, I'm home at seven every evening anyway, then I know I have my tablet. And for others this is very problematic. Then you have to make sure that you rather take something that has a higher threshold for developing resistance—I mean, when someone has a very, very irregular lifestyle. There really are some people who just party a lot and are out all weekend and so you just have to talk to people and see how reliably they can take the medication, to think together with them: What are the times in the day that you can keep, simply because compliance is infinitely important in HIV therapy" (Med1, r.202ff.).

The categorization of the interviewed physician conformed to the categorization of the DAH, an umbrella organization of regional AIDS groups in Germany, which also differentiated between four types of patient compliance. They emphasized that "this categorization is of course quite rough and must be seen absolutely neutral, i.e., everyone is who they are, and no one is better than the other. However, this rough categorization helps the doctor to choose the right HIV therapy" (DAH, 2005, p. 10). The implications of understanding patients in terms of belonging to different compliance categories were that physicians may choose a different form of therapy in dialogue with the patient depending on the anticipated compliance. Some treatment regimens must be taken reliably and precisely because missing out on pills would immediately decrease the effectiveness of the medication. Other therapy combinations are less sensitive in this regard and would still function if a dose is missed. The following categorization was published by the DAH (2005):

- The normal type: Is not 100% perfect but trying to take their medication regularly. They almost always succeed in doing this, with a few exceptions that are within the framework of normal human forgetfulness and indisposition. A few tools are usually sufficient to reduce the "normal" forgetfulness. This is the desired compliance-type.
- The holiday type: Usually take their medication reliably but not during vacation or at certain times of the year.
- The party type: Like to be on the go and may not have their medication with them and then skip one or two whole daily doses.
- The chaos type: Have no fixed daily routine and take their medication more or less regularly, depending on mood or external circumstances.

Compliance was assumed to depend on the lifestyle of the patient, their attitude to therapy, the frequency of medication intake and the side effects or interactions of the medication. Furthermore, the social environment was considered to play an important role (stigmatization or acceptance), as well as the doctor-patient relationship, the professional competence of the doctor, and the atmosphere and organization in the practice/clinic. Since taking the medication was largely beyond the control of a physician as this happened as part of the patient's daily routine, the practitioner had to be able to rely on the patient's statements: "And others are adamant that they really take everything every day and they get resistances where you just don't know—are they actually taking it that way, or are there other problems?" (Med1, r.271ff.). Compliance could not and should not be achieved through coercion and pressure (Bio2, r.147ff.). Rather, the patient had to be fully informed. The doctor-patient relationship was of crucial importance here. The patients had to be able to articulate their needs and wishes:

> "We do these analyses and try to find the safest possible combination of therapy for the patient. That is our concern, our job as virologists. And then the doctor's job—a doctor once said to me, 'Listen, have you ever considered that on the other end of that virus is a patient?' So, the patient has to want it, and do their part. And in the end, he has to take the medication—take it reliably. So, it's not enough to simply tell the patient, so, here's your prescription, bye. You have to get them to be invested" (Bio2, r.147ff.).

Recruiting patients to participating in their own care was an essential part of the job of the physicians, but it was not always simply a matter of encouraging compliance with a drug regimen. For example, while some HIV patients were characterized as "willing to take risks" and try out new therapies (Bio2, r.242ff.), some older patients were described as sticking to old treatment regimens out of habit or fear of changing something that works, despite simpler, e.g., single tablet regiments being available (Med1, r.469ff.). It was also lamented that many patients had lost their fear of the disease because HIV can now be treated well and has "lost its terror" (Bio2, r.47ff.).

In terms of good chances for therapy compliance, the single daily intake was regarded as the best available solution at the current time, even though long-acting implants were discussed by some of the interviewees. Compliance was conceptualized as the patient's submission to a specific daily routine: the necessity of eating at the same time every day and of taking responsibility for their own health through taking their medication regularly. This illustrates starkly that the role and responsibility attributed to patients was that they should do everything within their power to stay as healthy as possible.

The great importance of compliance also made patients powerful. In the end, therapy success was not only dependent on the choice and availability of the right (or tailored) therapy but crucially on patients reliably taking their medication. The members of the social worlds whom I interviewed anticipated that without the cooperation of the person with HIV, the therapy would not work. This implies a transfer of responsibility to the patients which was accompanied by morally charged discourses within the social worlds. For example, questions of patients' trustworthiness were raised, especially by the physician. Interestingly, there were situations where the logic of the field was disrupted. Physician Med1 (r.252ff.) spoke about a long-term patient who only later in the therapy confessed that for years they had only taken part of their medication. Still, the measurement of the viral load did not show a decline in therapy success, almost as if the medication had been taken as intended. Examples like this could function as disruptions in the social worlds and members then questioned their internal logic. This led to a recognition of the patient as a "black box" (Med1, r.268ff.) and to hypothesizing over which other medical reasons could lead to this unexpected outcome, such as metabolism.

Strikingly, though compliance was considered to be of paramount importance, the researched ML tools did not take this into account. At the core of the technology, resistance was conceptualized as being purely dependent on biological factors, while social factors, though constructed discursively as essential, were not included into the technology. This means that the assessment of the tailored therapy was always a multi-step process: viral properties were calculated by the technology and social features were taken into account by human experts (the physicians).

Categorization based on social determinants of health

Besides age, which was made relevant as feature of long-term HAART, being biologically female and in the reproductive age was a category raised by the physician and the virologists. This coincided with the guidelines outlining that from a clinical viewpoint, pregnancy is a special condition of an HIV infection and a specific antiretroviral treatment was prescribed (DAIG, 2020a). The desire to have children should be made transparent to the physician to avoid clinical complications. Virologist Bio2 described how research has not come far enough regarding antiretroviral therapy for pregnant women. According to HIV therapy guidelines, some active ingredients were either contraindicated or were not yet approved for pregnant patients due to a lack of information regarding potential teratogenic effects of the active ingredients. This made the range of therapy choices for pregnant people more narrow (DAIG, 2020b). The guidelines also recommended HIV tests at the beginning of all pregnancies to minimize the risk of newborn infection. Especially before giving birth, and later during lactation, the viral load should be low to avoid infecting the newborn (DAIG, 2020a). Physician Med1 and bioinformatician Bio2 (r.316ff.) made the category pregnancy relevant mostly at the intersection of "the migrant woman." This might be the case because, as Med1 (r.680ff.) stated, "German women" who tested positive for HIV were in the minority.

The category of migration background[24], especially regarding people of African descent, was made relevant by the physician and the virologists. The interviewees mentioned that patients' cultural backgrounds played a major role, as there were different understandings of values and living conditions. While the "enlightened" German patient (Bio2, r.334ff.) was said to be able to communicate with the practitioner without language barriers and could read and understand package inserts for medication, for the migrant women a "patient self-help project" was organized (Med1, r.54ff.). Within this project for HIV-positive people (mostly women) of African descent, expert patients were organized to function as translators during physician-patient visits and as general buddies/mentors about HIV. Migrant women were also more often than non-migrant women described as poor or as single parents with many kids (Med1, r.746ff.). Bio2 (r.160ff.) and the physician said that patients from other cultures, especially women from Africa, sometimes had different values, e.g., regarding childbirth and breastfeeding, and lifestyle habits, e.g., stronger roots in their own community, which were sometimes a source of conflict between patient and practitioner. In addition, Med1 and Bio2 spoke about communication problems that have an impact on compliance. Some of these factors were given as justification for establishing the patient self-help project (Med1, r.71ff.).

The way members of social worlds in this arena constructed this specific category at multiple intersections (migrant + woman + single mother + low income) could be regarded as special engagement and cultural sensitivity but also as *othering* or racist. Another reason given for the organization of the patient self-help project for HIV-positive people was what was described as the very common practice of gossiping within these communities.[25] This, it was observed, would discourage people from sharing their HIV status in the presence of a typical translator and they often felt safer with an expert patient translating for them (Med1, r.71ff.). From epidemiological studies it is known that for a long time, migration background was correlated with HIV infection (Robert Koch Institut [Robert Koch-Institute] [RKI], 2021). Now on the decline,

24 Migration background is a category specifically used in the German context since race is considered a problematic category after the Nazi regime and its use is generally avoided.

25 It is known that in the absence of professional(ized) interpreters (e.g, among new groups of immigrants to a place, where everybody knows everybody—and all their business), the code of ethical conduct which would otherwise ensure patient privacy can sometimes not yet have been established. This is justifiably experienced as gossip. However, this is a common feature of the early phases of any population's migration to a new country, regardless of origin, and not something specific to migrants from Africa.

people coming from Sub-Saharan Africa to Germany constituted a large proportion of people given initial HIV diagnoses in Germany between 2004 and 2021 (p. 7). The categories of men having sex with men and intravenous drug users, which were also present in the epidemiological study by the Robert Koch-Institute, were only indirectly addressed by the different interviewees.

In general, the categorization schemes seemed to be discursively constructed by members of the social worlds to create a simplified typology. These discourses were related to phenomena and materialities, and were used to "bring order" and "produce hard facts" ensuring an ability to act (Bosančić & Keller, 2019, p. 2). The importance of these orders can also be explained along the lines of Joan H. Fujimura's (1987, p. 257) framework, within which she aimed to conceptualize scientists' efforts to construct "doable problems." Doability was usually obtained through "the alignment of three levels of work organization: experiment, laboratory, and social world" (p. 261). Scientists achieved alignment by articulating—considering, collecting, coordinating, and integrating—tasks between these levels of work organizations (Fujimura, 1987). In my example, categorization schemes were important for the members of the social worlds in negotiating common understandings within and between the different social worlds and for being able to act in coordination to reach shared goals.

The productivity of the social world, social arena framework for analysis

The social world/arena framework could be used in many productive ways for this case study. First, with the help of the framework I showed who constitutes a social world, how social worlds are connected or not, which activities and goals are specific for a particular social world, and which ones they share. The social worlds/arenas map as the central tool of the analysis can also be adapted to the needs of the study. In this project, two iterations of the social worlds/arenas maps proved especially helpful. The map in Figure 2 depicts the social worlds and those actors not organized in social worlds as well as the way they are present between social worlds and in which categorization scheme they are discursively constructed. In the map in Figure 1 activities as well as specific and shared goals of the social worlds within the arena are depicted. Physicians and virologists worked together to find a suitable antiretroviral therapy and virologists and bioinformaticians, together with some physicians, shared in the development of the ML-based HIV TOS. In the two maps, I show how patients were simultaneously centered through the shared goal of the social worlds and marginalized in the way they were discursively constructed.

The analysis of specific and shared goals was important for the understanding of the longevity and success of the project and for a better understanding of the position of the patients especially in relation to their discursive construction by members of the social worlds in this arena. The specific shared goal in this arena was working together to ensure the survival of patients. Virologists were also interested in their scientific pursuit to understand the HI virus and bioinformaticians were invested in trying out their ML methods to optimize them and apply them to other (more promising) fields, e.g., cancer. The physician saw their role as treating patients. Under the umbrella goal of ensuring the survival of patients, the professions could pursue their own distinct aims. The specific and collective goals were based on the known reasoning within PM that more precise data gathered through/processed with algorithmic means were the link to a better understanding of a disease. In the end, the goal was to identify an optimized and individualized treatment, which it was thought would ensure the (longest) survival of patients.

Connecting the interviewees' goals to their views on the patients I can show that they mostly followed a logic accordant with their profession. For the bioinformaticians, patients were constructed as data because they needed data to optimize their tools. For the virologists, the virological information of the patients and their reaction to the treatment was key, as this shows how suitable the active ingredient suggestion based on resistance testing was. Physicians constructed the patients as people with HIV to be treated. In fact, people with HIV were simultaneously made beneficiary of the common goal and subject to the distinct goals of social worlds; though the practitioners in all the social worlds declared unanimously that their primary aim was to ensure patients' survival, it was clear that for members of two social worlds central to this arena, their own agenda was no less important.

A feminist theory of science was foundational for Clarke et al. (2015; see also Offenberger, 2019). Accordingly, she and her co-authors were particularly interested in identifying those who are at the margins of a social arena or who are marginalized by others. Individuals in an arena who are not connected with others in a similar situation—in analytical terms, organized in a social world—are more vulnerable. This can especially be true for the arena of healthcare. Using the framework, I have shown how HIV-positive people within the arena of HIV treatment were categorized by members of different social worlds. As has been established above, they were simultaneously central to the arena and marginalized because they were rarely constructed as actors acting in their own right. HIV-positive people seemed to be involved in the choice of their own treatment, but not in the optimization of the decision-making process or in the basic questions of whether further development of such tools was the right approach or of which other projects might be more beneficial for different patients.[26] One could say the categorization just followed the logic of the field and has no moral account. However, this can also be analyzed from a power-critical perspective. Virologists referring to and constructing people as "therapy failure" or "failure" raised the question of in which contexts members of the social world used these terms and in the presence of whom (Bio1, r.79, 446). What is the effect on the work of social worlds when their members construct people as failures? Which effects do discursive constructions have on the responsibilization of the patients and moralization of their behavior and compliance? In my analysis of these discursive constructions, I revealed the interests and value systems underpinning the way people with HIV are characterized and the nuances of marginalization which are evident across the social worlds. The question is whether people treated for HIV are benefiting enough, even when the goals of the other social worlds, i.e., learning more about the virus or optimizing a ML-based tool, are at the forefront.

I also demonstrated (Figure 2) that the physical presence of the implicated actors indeed changed how social worlds discursively construct them. The closer the members of a social world were to the patients, i.e., the more present physically people with HIV were with the practitioner, the more they were constructed by the practitioner as complex and human. This became apparent for example with the physician speaking at length about the influence of social determinants of health and emphasizing the need to decide on the treatment together with the patients. Physicians constructed the

[26] A shared decision making between patients and physicians/medical experts is an ideal. This project analyzed power asymmetries between experts and patients within the arena of personalized HIV therapy. Coming from a power-critical stance, the question of to which extent patients should be involved in such decision-making processes and in more far-reaching questions around development of such tools might be raised. Bearing in mind the responsibilization of patients in certain arenas, such as PM, I wondered whether a new conceptualization of patients as informed actors could pave the way for more empowering and participatory processes of decision-making also within the development of such tools or the necessity of other tools instead. In this I echo the suggestion of Tina Cook, Helen Atkin and Jane Wilcockson (2018) regarding participatory research for inclusive practice in the case of people with long term neurological conditions.

patients with different features: gender, ethnic background, age, and compliance type, which is connected to a type of lifestyle. These are general categories of people usually used within Western societies when referring to individuals as humans. However, the physician did not only refer to patients as humans but also constructed humans as patients in terms of compliance, entailing a specific role in this context where people are expected to adhere to specific activities to be considered worthy of help or are objects of examination. The virologists constructed HIV-positive people as objects to be examined with regard to resistant HI viruses and treated for them. Being experts for the viral part of HIV and its resistances, virologists were mostly interested in the outcome of the therapy. In the logic of this social world, the categorization of patients' therapy status was the most relevant one, because it was action-guiding. A "therapy-naive" person had still many choices for therapy, a "change-of therapy-patient" had to be checked for their treatment history (Bio1, r.446ff.). The occurrence of a "therapy-failure-patient" was the time when virologists could prove their skills to members of the other social worlds in the arena (Bio1, r.79). Work on these difficult cases also united actors in the arena in striving toward their shared goal of ensuring patient survival. Virologists—and the bioinformatician even more so—constructed patients predominantly as objects detached from their humanness: They were reduced to samples to be examined for resistant HI viruses and treated against HIV for the virologists and to carriers of data for the bioinformatician. It seemed that especially for the bioinformaticians, social determinants of health were not relevant.

In the arena of personalized HIV treatment, it was possible to rank the discursive constructions of HIV-positive people used by members of the social worlds from more human to more objectified in the order physicians, virologists, bioinformaticians. This coincides with the closeness of practitioners in the field to the patients. The less members of a social world centered people as humans within their work, the more objectifying the discursive construction of people with HIV was. It is important, however, that commonness not breed complacency, because the consequences of this objectification can be grave for people seeking treatment for HIV: As practitioners in the fields of medicine and healthcare embrace more data-driven applications such as PM, patients' risk being inadequately understood as complex and human, which, in the worst case, can compromise their care. Patients might be reduced to their data points and dehumanized. This can represent an even bigger problem for marginalized people, whose discrimination stemming from systemic inequalities might then be dismissed as individual problems (Baumgartner, 2021b; Prainsack, 2017b).

In this case study, the physicians who were closest to the patients did not work directly with the tools. Virologists and bioinformaticians developed the ML tool, the virologists used the tool and conveyed the outcome to the physicians. The physicians then decided on the suitable antiretroviral medication together with the patient. They only provided the information about the therapy outcome back to the virologists, so the virologists could validate their assessments and the quality of the tool's outcome, similar to a feedback loop. It is difficult to retrospectively analyze how the introduction of algorithmic HIV TOS might have changed the way treated people were conceived of by practitioners in the field of HIV treatment. One could surmise that the way of constructing patients as carriers of data—a construction introduced by the bioinformaticians with the beginning of ML-based HIV TOS—would circulate in the whole arena. However, at least the interviewed physician, who was not part of the development, did not construct their patients this way (while at the same time speaking of other ML-based applications and presenting knowledge about the importance of patient data for these). This point would have to be explored with physicians who were part of the development in order to understand whether more involvement with PM changes the way they discursively construct patients.

Through the analysis we can also see how closely members of different worlds aligned in their understanding of the patients. In my case study, some parts of the categorization schemes were present throughout the social worlds (e.g., "therapy failures"), while categorization based on social determinants of health were shared by physicians and virologists, and the categorization scheme based on compliance types and lifestyles was used only by the physician. Shared discursive constructions might be interpreted as closer relationships between actors in social worlds because a shared understanding and ways of thinking about certain subjects might be indicated through them. A categorization scheme used by just one social world could mean that its logic is only relevant for this very social world, such as thinking in lifestyle categories to anticipate compliance was mostly relevant for physicians, who based their decision on the tailored therapy on this assessment. Interestingly, this important information was not adopted by bioinformaticians, who could have introduced it in their tool. This hints at the way HIV infection was constructed for the work in different social worlds. Seemingly, bioinformaticians above all constructed it as a purely biological disease, not conceptualizing it as also being influenced by social factors. The question is whether this way of thinking migrates through the ML tool back to the users of the tool (in this case virologists and not physicians).

In this explorative study I show how productive SA can be when analyzing social worlds, arenas, and implicated actors. The framework is suitable for researching case studies of digitization in medicine and healthcare such as PM and introduction of AI in different fields. The social worlds/arenas maps were also used to depict different points of time within the arena, e.g., pre-AI introduction and post AI-introduction. With its strong focus on analyzing power asymmetries and focusing on marginalized actors/actants, SA is particularly suitable for a critical analysis of newly introduced AI in different fields. SA researchers can flesh out complexities in the social arenas during analysis. Tinkering with and adapting social worlds/arenas maps fosters the creative process and the development of an understanding of the analyzed situation. With SA, researchers can distinguish between specific and shared goals and their influence on members of social worlds as well as on actors in the arena who are not collectively organized. Researchers can also analyze logics and discursive constructions within different social worlds and make sense of similarities and differences between the worlds. The analysis reported here led me to important findings, such as the correlation between the degree of closeness of implicated actors to members of social worlds and the respective construction of those implicated actors as more/less complex and human. Through my mapping, I was able to show that centering actors in the arena does not prevent the same actors from being marginalized, as seen in the various discursive constructions discussed above. Thus, with the help of SA framework, marginalization could be conceptualized as a nuanced process that can be analyzed as it happens across social worlds.

Outlook

From the current study it is evident that people treated against HIV, despite being at the center of the arena—literally giving the arena of concern its raison d'etre, were simultaneously implicated by others and marginalized. Members of professional groups whose specific professional goals are foregrounded in the arena should take care that this does not result in neglect of patients' needs. The umbrella goal of ensuring survival of the patients would have to be made more concrete and filled with patients' perspectives in addition to the specific goals of the social worlds involved so far: What do HIV-positive people need nowadays? How might already existing support groups, such as DAH, participate in representing and formulating those needs during the higher-level decision-making processes regarding suitable therapies and within the development of new tools? Currently, with single tablet regimens the survival of most HIV-positive people can be ensured without complications. Is the development of new and more precise, e.g., NGS-based HIV TOS, which might increase digitization and a deeper analysis of patient data, the right way to go? Or might other activities which put patients' health(care) needs at the center be more beneficial for the health and well-being of HIV-positive people? I advocate for participatory projects which include all stakeholders and a diversity of HIV-positive people, to collectively approach the question of which HIV TOS developments should be pursued.

Acknowledgments

I would like to thank Darja Burljaev for her help in early stages of the paper and Tamara Schwertel and two anonymous reviewers for their insightful feedback on the paper, as well as Sarah B. Evans-Jordan, Ursula Offenberger, Birte Kimmerle and participants of the working conference (especially Marc Bubeck and Bianca Jansky) and members of the Tübinger SA interpretation group for their comments on earlier versions of this manuscript. I also thank Aljosha Victor Kannewurf, Fanni Weber and Sebastian Bartelheim for assistance in editing and proofreading the article.

3.3 Precision medicine for HIV therapy: A tale of successful risk and uncertainty management?

Introduction

Precision medicine (PM) is, since the early 2000s, one of the dominant paradigms in medicine besides evidence-based medicine (EBM). Both strive to optimize healthcare, e.g., taking into account the newest knowledge and thereby making medicine safer (McCoy, 2020). All in all, the objective is to improve health for patients. Both rationales rely on knowledge with which danger, risk and uncertainty can be managed. HIV and AIDS have at first been a tremendous challenge for society as well as for medicine. The fear of HIV being a pandemic could only be appeased after the detection of the HI virus could lead to the development of efficient medication in the mid 1990s. However, even with the availability of highly active antiretroviral therapy, the danger of the ever-mutating virus to escape its treatment was a death risk for HIV-positive people who would run out of available treatment options (Epstein, 1996). More than 15 years later and with the advent of efficient single tablet regimes, HIV is considered a chronic disease and its history is framed as a big success story of medicine and science. Nowadays, antiretroviral treatment (ART) in most cases efficiently contains the danger of dying from HIV/AIDS. Nevertheless, about 10% of HIV positive people struggle with resistant forms of HIV (Bio01). The search for suitable treatment within HIV with HIV treatment optimization tools (HIV TOS) involves precision medicine and is one of the first successful applications of this medical paradigm. This paper provides the analysis of a case study of HIV TOS developed in Germany starting from the early 2000s and used until today. Tools and practices surrounding them will be analyzed applying theories on risk and uncertainty from (medical) sociology. Risk, uncertainty, and certainty formulated by the stakeholders of the field are interpreted from a social constructivist perspective. The mitigation of risk and uncertainty is analyzed on an individual, an organizational, and a technical level. It will be shown how the profession of virologists constructs their own professional relevance based on the construction of certainty and uncertainty during the process of coming to treatment suggestions for HIV positive people. Lastly there will be a discussion of how this plays out in precision medicine, which is used within evidence-based medicine as dominant paradigm.

HIV/AIDS and precision medicine

The term AIDS for acquired immunodeficiency syndrome was applied soon after the first young men in the early 1980s died of a very specific form of pneumonia. Western Societies of the 1980s were fear-driven because of this disease and progress could only be made once the HI virus was detected as its cause. Researchers found out about viral transmission by blood and about the viral surviving strategy: mutation. The virus mutated so fast that in these early times, HIV positive people ran out of medication quickly, due to multiple resistances to the then available active ingredients. Only after a while, virologists would understand how to combine different active ingredients into an effective antiretroviral treatment (Epstein, 1996). Virologists also realized that they had to track the specific virus cocktail within a person very closely, in order to choose the right medication the virus was susceptible to and which would not be evaded by mutations (Beerenwinkel et al., 2001). This rationale, finding the most suitable therapy for the patient in a very specific health situation is a key feature of PM (Baumgartner, 2021b; Prainsack, 2017a).

The available knowledge about the viral mutations was first organized in lookup tables which constituted an archive for HIV resistances in relation to active ingredients. After a while, this knowledge was digitalized in the form of rules into algorithmic tools that were made available online. With the advent of machine learning (ML), the first tools were developed that would suggest active ingredients or treatment regimens based on viral data input (Beerenwinkel et al., 2001). They would function at the same time as clinical decision support tool (based on PM). This paper will analyze these algorithmic HIV treatment optimization tools regarding the way uncertainty is mitigated and certainty is reached with their help as they manage but produce risk.

Material and methods

The case study is based on expert interviews with one physician, an infectiologist and expert on HIV treatments, two virologists (Bio01 and Bio02, one interviewed twice) and two bioinformaticians (DS06 and DS07). All interviewees except for the infectiologist have been involved in the development of the analyzed HIV TOS. Additionally, publications of the development team and others on the applications (as cited on the respective homepages of the tools) were analyzed. Interviews were transcribed verbatim and coded according to Grounded Theory Methodology (Strauss & Corbin, 1994).

The development and usage of the ML-based tool was also analyzed with situational analysis according to Adele Clarke and scholars (Clarke et al., 2018). More details about the analysis can be found in a previous publication (Baumgartner, 2023a). Early on, it became evident that many practices the field had established within the process of therapy suggestion had the goal to mitigate risk and uncertainty. The following section introduces used theories from the sociology of risk and uncertainty and from medical sociology. I will also explain how the management of risk and uncertainty is related to knowledge and provide a framework for the analysis of the role of (un)certainty in medicine and healthcare.

Sociology of risk and uncertainty

Risk can be understood as "a material or symbolic danger or harm, or an alleged negative future event" which from a technical perspective is calculated as "probability and extent of an (undesired) event" (Zinn, 2008, pp. 4, 173). Depending on the discipline and theory risk can be conceptualized as objectively existent or as "being socially mediated or even socially constructed independent of its objective existence" (Zinn, 2008, p. 4). From the different approaches, the realist perspective deals with events or dangers, which are thought to be assessed objectively and calculated as the probability of an event multiplied with the damage it produces. Probabilistic risk calculations are an example of these and are also a way to deal with an uncertain future (Porter, 2020; Zinn, 2008). Based on this conceptualization of risk, the related uncertainty is seen as lack of knowledge which can be overcome with additional objective knowledge produced by science and more rigorous analyses (Zinn, 2008). Put differently, within certain domains such as science and medicine, more knowledge can facilitate the notion of reaching more certainty. While technical-science disciplines mostly regard risks from a realist perspective, most approaches within sociology and science and technology studies are rather interested in the conceptualization of risks as socially constructed no matter if measurably present or not (Zinn, 2008).
Uncertainty is also said to be the "relative degree of our inability to predict the future" (Seely, 2013, p. 67). It becomes especially relevant once the danger is expected to lead to an undesired event which might have a significant effect on life, such as a wrongly or untreated HIV infection would first lead to HI virus mutations and eventually to the death of the infected person (Epstein, 1996; Zinn, 2016). While this paper

acknowledges that danger, risk, and uncertainty can be objectively present, the analysis predominantly deals with how actors in the field have socially constructed them. Whenever both certainty and uncertainty are meant, I will use the term "(un)certainty". Zinn speaks of different strategies to manage risk and uncertainty. Rational strategies involve calculation or weighing pros and cons, in-between strategies involve intuition and trust, and not-rational strategies involve hope and belief. These show that for managing uncertainty, apart from data and knowledge, also intuition and trust are used (Schulz & Zinn, 2023; Zinn, 2016). The presence of experts is a main source of trust (Schulz & Zinn, 2023). However, in some fields trust in experts does not suffice and there is hope that more certainty can be reached by introducing digital technologies, e.g., to make all relevant knowledge available or create new knowledge. These technologies again pose new problems of trust (Jermutus et al., 2022). Thus, with the rise of algorithmic or even more so with ML-based tools new questions about trust and technology have been raised and a whole field of research is interested in asking how and if we (can) trust technology such as clinical decision support systems (CDSS) (Baumgartner et al., 2023; European Commission, Directorate-General for Communications Networks, Content and Technology, 2020; Yang et al., 2016).

A sociological categorization of knowledge for precision medicine, evidence-based medicine and beyond

Different types of knowledge play an important role when it comes to managing (un)certainty. Certain aspects of the following sociological conceptualization of knowledge types will remind us also of Zinn's risk mitigation strategies. Rammert (2000, p. 3) drawing on Polanyi (2009, p. 22) distinguishes between "explicit knowledge", "not-explicit knowledge" and "implicit knowledge" ("tacit knowledge" according to Polanyi). Explicit knowledge according to Polanyi can e.g., stem from science. According to Rammert's sociological reconceptualization, explicit knowledge can also be knowledge which represents a rational choice, concerns formulated rules or is part of an algorithmic model. Not-explicit knowledge, according to Rammert, can be routines, common or informal knowledge, or knowledge from socialization, e.g., coming from professional experience. Implicit knowledge according to Polanyi is, e.g., intuition.

Medicine has a long history of debate which knowledge is trustworthy enough as base of medical decisions and there exist elaborated sociological analysis thereof (Parsons, 2005; Schubert, 2022). In Parsons' (2005 (1951)) ground-laying text on the structure and function of modern medicine, he also discussed uncertainties within medical diagnosis and framed diagnosis as ideal typical rational process on the one hand, while on the other hand admitting that the messier praxis was more difficult to analyze. The question which type of knowledge diagnosis would be based on was from there on at the center of sociological analysis (Schubert, 2022).
There are different movements within medicine invested in developing procedures based on scientific knowledge. EBM is a "self-proclaimed movement" within medicine and relevant "paradigm" since the early 1990s , p. 1001). It revolves around "quality of evidence" considering scientific knowledge as gold standard with a clear hierarchy of evidence seeing randomized clinical trials (RCTs) at the top (McCoy, 2020; Nardini et al., 2012, p. 1001). EBM has an empiricist logic and is epistemologically concerned with standardization based on high quality data from systematic and aggregated analyses of evidence based on large patient groups (McCoy, 2020). PM's presence within medicine stems from the early 2000s and needed the support of emerging technologies within bioinformatics and genomics (Au, 2021; McCoy, 2020; Nardini et al., 2012). PM aims at finding individualized treatment for patients based on subgroup analyses of small groups of patients and patient-specific data. PM's rationalist logic is concerned with "the importance of mechanistic physiology theory in contribution to the body of medical evidence" (McCoy, 2020, p. 28). This means that both PM and EBM are invested in basing decisions on scientific knowledge. Even if many stakeholders in the field are invested in these two paradigms it is known that within clinical routine experiential knowledge and routine is more important than the field itself would acknowledge (Schubert, 2007).

(Un)certainty in science and medicine

Both science and medicine seek to manage uncertainty and the risks associated with it. Epidemiology and toxicology, e.g., are dedicated to calculations of risk and uncertainty (Ghosh, 2004). One of the main goals of scientific progress is to reduce uncertainty by producing more knowledge. "Diagnostic and treatment decisions are fundamentally concerned with the application of medical knowledge amidst uncertainty" (Fox, 1980; McCoy, 2020, p. 27). Different movements such as EBM and PM provide frameworks within medicine to guide healthcare professionals and healthcare organizations in diagnosis and treatment-related decisions. Healthcare professionals provide guidance to lay people when assisting them to navigate uncertainties within their health status and treatment thereof (Ghosh, 2004). Lay people trust healthcare experts also because of their assumed expert knowledge (Bogner et al., 2009). However, also these experts must minimize uncertainty. This is especially true in medicine and healthcare where knowledge about disease, diagnosis and therapy has been growing considerably within the last decades and there are challenges to manage this knowledge and to know what can be known and what cannot be known (Fox, 1980; Furlow, 2020; Horn, 2001). This need to manage uncertainty has also increased since patients are more educated, tolerate less uncertainty and are more likely to take legal action in cases of misdiagnosis and mistreatment (Fox, 1980; Seely, 2013). In an age of increased responsibilization of individuals as well as organizations there are measures in place to manage risks and uncertainties on an individual and on an organizational level (Hood & Rothstein, 2001; Zinn, 2016). The goal of managing risk and uncertainty – among other things with more knowledge – should eventually lead to more certainty and to better health.

The question if uncertainty and/or certainty is a key feature of medicine has been a contested topic within sociology of medicine (Atkinson, 1984). Medical sociologist Fox (1980) achieved fame for her ethnographic studies on how medical students master uncertainty in medicine as part of their socialization to medical professionals. Fox (1957, pp. 208-9) focused in her classification on medical knowledge and distinguished between "incomplete or imperfect mastery of available knowledge", "limitation in current medical knowledge" and failing to distinguish between the former two. Within their socialization as medical professionals trainees learn to control different kinds of uncertainties (Light, 1979). Light (1979, pp. 311-313) expanded Fox' concept of uncertainty in medicine with additional "kinds of uncertainty", such as uncertainty in diagnosis, - treatment, and - "client's response". Ways to control these uncertainties are mastering knowledge and gaining professional experience (Light, 1979). Seely's (2013, p. 68) more recent classification of uncertainty within present day healthcare points out "uncertainty in basic science", -"in health-care practice", -"in health-care management" and -"in patient-physician communication" . He also offers a classification of uncertainty related to knowledge in healthcare with some similarities to Fox' classification where uncertainty is divided into "informational uncertainty" and "intrinsic uncertainty" (Seely, 2013, p. 67). The former deals with what can be known and how precise this knowledge is being close to Fox' mastery of the available knowledge. Informational uncertainty can be reduced with a growing body of knowledge. In contrast, intrinsic uncertainty is close to Fox's concept of uncertainty of the professional knowledge itself. This type of uncertainty cannot be reduced with additional knowledge but uncertainty will stay the same (Seely, 2013). Atkinson (1984) voiced critique regarding Fox' approach of putting uncertainty alone as central aspect of medical (and scientific) formation and key feature of medicine. According to him, uncertainty must be seen in a more differentiated way and certainty must be added as concept. "Both 'certainty' and 'uncertainty' should equally be seen as features of the social construction and definition of medical discourse" (Atkinson, 198, p. 954). An important question for him is when which aspect gets constructed by scientists and medical professionals. Medical knowledge can be constructed as uncertain, but focusing on professional judgement and relying on medical dogmatism certainty can be constructed. He bases his argument on Schütz' phenomenology and adds that experts also see the world as lay people. In this role they trust their common sense, their experience and routine and feel certain about the world. Similarly to healthcare professionals, also scientists have to manage uncertainty. Star (1985) showed how scientists within medicine manage uncertainty within their local settings and their disciplines.

This case study is concerned with virologists who are experts within the field of HIV and its treatment. They are not per se medical professionals. However, they are very closely involved in medical decisions around treatment choice within ART. As Light (1979, p. 312) writes: "Thus uncertainties of treatment threaten the *raison d'étre* of a profession and must be controlled". Virologists play a major role in the process of treatment selection which means they are invested in doing their job right to protect their profession. I will argue that this aspect plays a major role in the virologists' construction of (un)certainty and the relevance of their profession. Additionally, the theorizing of medical sociology goes well beyond medicine and healthcare, even more so since the borders between medicine and medical science seem more and more blurred (Clarke et al., 2003).

Medical and scientific experts are not the only ones able to provide more certainty within healthcare. With digitalization, ML and artificial intelligence, a new hope has emerged to manage uncertainty with these tools in biomedical research and healthcare: Either by producing new knowledge based on different workings of these algorithmic tools which can, e.g., be used to find patterns in large amounts of data that humans would not be able to find, or with calculating probabilities and precise predictions based on a higher volume of and more fine-grained data. With these tools, predictions of what might happen in the future, being one of the central possibilities to reduce uncertainty, suddenly seems to be a viable option. (Baumgartner et al., 2023; Schubert, 2022)

Personalized HIV therapy and HIV TOS: Development and function within the field

As the knowledge about HIV grew, virologists as experts for the HI virus together with infectiologists as experts for its treatment worked together identifying efficient and suitable therapies for HIV positive people based on containing drug resistances of the virus (Bio01). For their shared assessment they first needed to identify the type of viruses present in the infected person which was assessed during regular checkups based on genotypic drug-resistance tests done in labs. In 2006 Liu and Shafer (2006, p. 1) described the selection of a suitable ART as complex challenge were algorithmic HIV TOS were regarded as necessary for experts to interpret the results of these tests because there are i.a., “complex interactions among the many mutations that contribute to drug resistance”. The tools provide information on the “clinical significance of mutations” and thus, “provide clinicians with data that help them to make” the “most-informed treatment decisions for their patients” (Liu and Shafer, 2006, pp. 6-7). Virologists suggest an individualized treatment based on the outcome of the HIV TOS and their own knowledge about the HI virus’ susceptibilities to active ingredients. Physicians weigh in their knowledge about the patient (e.g., on concomitant diseases and compliance) and consolidate with the latter their final suggestion (Bio01). Over the years, different HIV treatment systems have been developed both in the public and private sector (Liu & Shafer, 2006). In this paper, I will focus on the currently used algorithmic systems, namely ML-based genotype-phenotype systems and rules-based systems. Firstly, I will provide a short historical overview of the different systems and show the function they have in the field. Secondly, the usage of the tools by virologists will be outlined. Both aspects will be analyzed applying the sociology of risk and uncertainty which helps to explain the reason the tools were developed in the first place and the way the tools are being used.

The development of HIV TOS started with analog lookup tables where experts explicated and organized the available knowledge about known resistances of different HI viruses against active ingredients (DS07) (Beerenwinkel et al., 2001). As digitization progressed this scientific knowledge was programmed in form of algorithmic rules into HIV TOS which would reproduce the knowledge in their outcomes. The tools which the experts in this case study have been using are HIV-Grade and HIVdb (Bio01). They were both developed around 2000: HIV-Grade by a German organization of virologists – part of them my interviewees – and HIVdb by researchers of the university of Stanford (Bio02) (Liu & Shafer, 2006; Rhee, 2003; Shafer, 2006). Beerenwinkel et al. (2003, p. i17) describe these as "rule-based expert systems designed for finding optimal resistance-avoiding combination therapies" which aim at "knowledge-based therapy optimization". They were developed extracting classification rules from scientific lit. E.g., also with data made available from companies after the approval of a new antiretroviral substance (Bio01) (Zazzi et al., 2016). They were updated with new knowledge as available. From a sociological perspective the rule-based tools can also be regarded as knowledge-based decision support systems (Ozaydin et al., 2016). Within the social world, they were given the role of an archive or knowledge repository (Horsley, 2021). From an organizational perspective, HIV-Grade is a community effort of virologists to have a consolidated knowledge base and to profit collectively from the newest scientific and/or validated knowledge. One could also say they implemented a common standard of knowledge by sharing knowledge in the form of digitalized rules.

In the early 2000s, a cooperation between bioinformaticians, virologists and clinicians started developing a family of ML-based HIV TOS. The first developed tool being also one of the first tools applied within PM which is still used today (Bio01). The virologists had data from phenotype assays from cases where available therapies would fail to work. This first "geno-to-pheno-pairs" were obtained from blood samples of these patients (cf. Bio 01). Once approached by bioinformaticians interested in collaborations virologists and bioinformaticians realized they could use this data as training data[27] and saw the opportunity to create new knowledge with the help of these ML-based tools (DS06, DS07, Bio02). Based on the training data, the ML algorithms

[27] Means genotypes of the HI virus were exposed in lab cell-culture to antiretroviral ingredients observing the reaction, i.e., obtaining phenotypes of how the virus would be eliminated by the active ingredient or be resistant to it. (Bio02)

would calculate resistance factors which are the basis of the assessment of resistance of a known HI virus to a known active ingredient. In the end, the experts aimed to predict the phenotype - how the virus will react to an ingredient - from a known genotype regarding different available antiretroviral ingredients. Data input during usage of the tools is the human HI-viral RNA derived from a blood sample of a HIV-positive person. Output is the probability calculation about which active ingredient this specific HI virus[28] might be susceptible to or resistant to (DS06, Bio02) (Beerenwinkel et al., 2002). From the beginning, the idea was that knowledge should be generated in the form of newly found correlations in form of resistance factors. From a sociological perspective, these newer tools can also be regarded as data-based tools with a common goal of producing new knowledge. The training data for these early forms of the ML-tool has to be produced in labs and can thus be regarded as explicit or scientific knowledge. Therefore, these ML-based tools demand a lot of work during development, maintenance, and optimization. In contrast, once a new active compound is approved for rule-based TOS, scientific literature might be made available by the company and can be entered. However, the different types of tools are also connected to each other when it comes to the produced knowledge. Beerenwinkel et al. (2002, p. 8274) described the advent of ML-based tools as follows: "We generated decision trees that identify patterns of several positions predictive of drug resistance or susceptibility. The decision tree method appears adequate because the classification knowledge is presented in a form that human experts can easily understand and examine, and because it is capable of representing effects of interactions between different mutations. From decision trees it is easy to derive rules, the currently dominating form of representing HIV-1 resistance knowledge." This early assessment shows that, already at the beginning, the tools were developed in a way that knowledge could be transferred from ML-based tools to rules-based tools. In the early times of this decision support system (DSS) development, this was done to have enough information to address critique, later to make new knowledge generated by ML-based tools available to rules-based tools (DS07). I will lay out the transfer of knowledge from one tool to the other in a later section. All described digital tools can be openly accessed through their homepages. Thus, they are available to everyone with an internet connection who has the

[28] It is more a cocktail of HI viruses. However, for the sake of simplification I will speak about "the HI virus".

input data (HI-viral RNA sequences) to obtain data output. None of the described tools is approved as a medical device (Bio01).

Personalized therapy with HIV TOS from the perspective of sociology of risk and uncertainty

What does danger, risk and (un)certainty constitute within HIV treatment? The danger for a HIV positive person is first and foremost to die from a HIV-related illness, such as AIDS because the antiretroviral therapy would not work sufficiently. Since HIV is a rapidly mutating virus which can easily escape ART, management of drug resistance is key. Knowledge about drug resistance which determines the effectiveness of ART is therefore regarded as key for virologists to decide which treatment to suggest. Virologists deliver this treatment suggestion to the physician who decides together with the patient about the treatment. With their involvement in the suggestion and decision virologists and physicians also take a risk, on an individual professional level as well as on an organizational level. As virologists decide about their treatment suggestion, they must manage the uncertainties that come with the adopted risk. This is done in manifold ways as the analysis will show.

Knowledge is an important way within science and medicine to manage uncertainty. E.g., knowledge about different HI viruses and knowledge about approved active ingredients is crucial for experts to decide about a suitable and efficient ART. They also need to know how the active ingredients work, which viral subtype the ingredients can suppress and in which combination they work best. As we will see later in a transcript the decision about the therapy suggestion is based on different types of knowledge such as scientific and experiential knowledge. However, scientific, or explicit knowledge can easier be managed by digital DSS and will therefore be the first focus of the paper. Since the discovery of the HI virus and the first antiretroviral ingredient, scientific knowledge about ART increased and so did knowledge about the susceptibility of different HI viral subtypes to different active ingredients. This also means that the risk of not considering relevant knowledge for a therapy suggestion increased and with it the perception of the risk that this would wrongly influence the choice of therapy and thereby endanger the patient's health.

In addition to being based on the most accurate and recent knowledge, the treatment decision must also be accessible. Initially, lookup tables were created, which should help experts by providing available knowledge to be taken into consideration for their informed decision. This could be regarded as reduction of complexity and at the same time as managing the risk of forgetting important knowledge. It could be a pragmatic decision as well as an indication for not trusting experts enough to have all the available knowledge organized individually. Additionally, the development of lookup tables can be considered a step to make knowledge available globally and thereby was a democratization of knowledge. This step was also a form of standardizing quality of care which was important within EBM as dominant paradigm within medicine. Following technological development, it seemed safer to transfer the knowledge from lookup tables to algorithmic knowledge-based systems such as HIV Grade and HIVdb. This seemed to be a logical step as digitization of medicine took place and would also make the knowledge available more widely. A process of precaution was established to ensure that the right rules were part of this knowledge-based system: a panel of experts would decide which rule to introduce and which not (Bio01). This means human experts have been deciding which knowledge is regarded reliable (or scientific) enough to be introduced as rules in these tools.

However, with the advent of ML, current scientific knowledge did not seem to be enough. Suddenly, it seemed possible to find new knowledge – in the form of predictions of resistance factors – with the help of ML-based tools. Once this option was available, the field operating in the mindset of offering the best possible treatment based on the best available knowledge, which are important in both PM and EBM, this possibility could hardly be skipped. Also, from a pragmatist perspective, predictions are often applied to solve relevant but uncertain future developments (Schubert, 2015). The perspective that the field has on risk can be interpreted as realist considering the "production of more objective knowledge" as "superior way to manage respective uncertainties" (Zinn, 2008, p. 5). However, the experts within this field seemed to know that the correlations generated by the ML-tool cannot simply be interpreted as causalities. They see the need to make this knowledge scientific. This is again done through a panel of experts. This connection between the two different types of tools was programmed into the ML-based tool. At the beginning of its development, the team saw the risk that other experts would not trust the outcomes of the ML-based tools. They took measures for that by choosing a statistical model that is transparent (first decision trees, then support vector machines) and provides the possibility to extract newly generated rules from it. In this way, the new rules can be checked, transferred to other rule-based tools and this knew knowledge can be made available at the same time as the rule-based tools is updated with the newest knowledge. In practice, when new data has been generated or a new rule is found, a panel of experts decides if the rule is plausible, which happens twice a year. This can be interpreted as a validation process which functions to mitigate the risk of introducing incorrect rules. Only if they agree that the information is trustworthy it will be integrated as new rule in the rules-based tool. This means that knowledge of ML-based systems is made scientific (or turned into "objective knowledge") by experts with the help of the established scientific knowledge and expert knowledge. Therefore, in the end, rule-based systems do not only incorporate scientific knowledge but also knowledge consolidated as scientific (enough) to be part of the tool.

Usage of HIV TOS

While the previous historic outline might suggest that this is a story of historical sequence in practice some of the tools are used interchangeably or side by side. Virologist Bio01 who was part of the development of the tools describes the usage as follows:

> *"We use these three. The German HIV-Grade, [name of ML-based tool], and HIVdb. // Ah, yeah. Crass. This means, you practically put similar data into all three...// Yes, we have this function in HIV-Grade, when you put in the sequence, you can check which databases you want to query, too. And then the prediction is provided for all the ones you want to have, in comparison right next to each other. That's why then it's so much easier. But you don't always have the newest version of HIVdb, so sometimes, when we have doubts or so, we access HIVdb directly and then test the sequence again and get the information."*

This virologist describes that rules-based tools HIV-Grade and HIVdb are used side by side with the ML-based tool. This practice is so common that the possibility is provided as option within the tools to obtain predictions from multiple tools, i.e., comparison of the outcome from different tools is encouraged. To be on the safe side, sometimes even the newer version of HIVdb is accessed separately. The parallel usage of the tool which is also programmed onto the homepage of the ML-tool suggests that the people who developed it did not trust the outcome of one tool alone or they considered it too risky just to use one tool without checking it. Literature also discusses that different systems can produce different interpretations (Liu & Shafer, 2006). When asked how they go about it if the outcome between the tools differ, virologist Bio01 responds they would then use the newest information and/or the worst-case.

In addition to the usage of three different tools, virologist Bio01 stresses their own responsibility in the process of active ingredient suggestion. The outcomes of the tools cannot be trusted alone but the virologists must keep themselves scientifically up to date and not only have to check the outcome of the DSS for plausibility but also need to be cognisant of the available scientific data and able to interpret the scientific data regarding its scientific robustness:

> *"Therefore, I say, so the resistance finding ultimately, is not just these predictions, but also our, our information situation. We also visit congresses and read papers and so on and keep ourselves up to date also in these regards. // The own plausibility check, so to speak.// Right! Yes, but this must be, because you always have to be up-to-date. And you can't reach that with any web page. So even with a rule-based [tool], the experts must meet and reach a consensus regarding a mutation. This is not always achievable on a daily basis, but if you were at a congress last week and the newest data on an active ingredient was presented and you know better at that moment, then that has to be included. I mean, you don't want to withhold that from the patient just because you say, no, but this website says. You don't do that, of course. // Yes, yes. And in case of doubt, what would one do? // In case of doubt? You mean, if the [web]site says something different? // Yes. // As I said, I would rather [use] the newest, the newest findings. Which you then know yourself. But of course, depending on whether you find them relevant or not. So that too, of course."*

This is also corroborated by literature. Liu and Shafer (2006) write that because of several shortcomings (e.g. the HIV TOS does not take into account previous drug-resistance test results, CD4 cell counts, ART history, and antiretroviral cross-resistances) the outcome is only a guidance for the ultimate decision. With the transcript in mind, one could also conclude that the patient must be treated with the latest scientific knowledge in mind. This reminds of the values of EBM and – if the claim to understand what happens mechanistically is added – also of the values of PM. Applying Seely's framework, we can conclude that they strive to reduce informational uncertainty. Speaking with Fox and Light, the available knowledge must be mastered and there has to be a frame for enabling "as much cognitive command of the situation as possible" (Fox, 1980, p. 7). Looking at the aspect of democratizing knowledge, the ideal is that the knowledge of the expert must be at least as up to date as the knowledge of the profession.

Even if the technology might have been introduced to manage uncertainty, we have just learned that none of the technologies can stand alone. Experts must control the tools and their suggestions and manage the uncertainties that arise with their application. In the following section, I will outline the different strategies they take on to mitigate uncertainty when using algorithmic HIV TOS.[29]

[29] During development of the tools risk mitigation strategies are applied, too. These involve the following strategies: Checking data quality and choice of training data (for ML-based tools). Before the tool is made freely available to the public, it is validated by experts (bioinformaticians, computer scientists, virologists).

Table 1: Risk mitigating strategies during decision of ART suggestion with HIV TOS

Risk mitigating strategy	Level	Description/example of risk mitigating strategy
Usage of HIV TOS	organizational/individual	To safeguard a treatment decision HIV TOS are used
Parallel HIV TOS usage	technical/individual	Offered on homepage and entered individually in different tools
Plausibility check	individual(/organizational)	Done by virologists using consolidated scientific knowledge and experiential knowledge
Individual practices during decision taking	individual	E.g., virologists using worst case scenario in a difficult decision
Practices of safeguarding	organizational/individual	Documentation of process leading to decision, for transparency and legal protection of the individual expert and the organization.
Divided responsibility of decision making	organizational	Virologist as expert for viral aspects of HIV (deciding based on their own knowledge and the information from HIV TOS), physician as expert for patient. In the end, physician decides together with patient on the medication. The decision is informed by the information provided by the virologist.

As the choice of active ingredients is, in the end, the experts' choice, also experiential knowledge of the experts can come into play and is formulated as key form of knowledge by virologist Bio01:

> *"So, the findings that we then create, here in the virology in [name omitted] and I think also in most virological institutes, it is so that ultimately, we also let the experience, our own experience also flow in. Depending on what experience we have made, we say, okay, no, here we give priority to this system and here to this system rather, yes, right. It's a bit difficult of course and sometimes it's also emotion-driven."*

This transcript shows us that for the treatment suggestion scientific knowledge and experiential knowledge including intuition are used. [30] According to Polanyi (2009) and Rammert (2000) it would mean explicit knowledge, not-explicit knowledge and implicit knowledge are used to come to a decision. Interpreting the knowledge as "decision-making strategy" according to Zinn (2016: 348) virologists use "rational"

[30] This part of making the own experiential knowledge so strong could also be interpreted as experts underlining the importance of their role in the whole process after the introduction of algorithmic tools in the field.

(calculation, risk benefit ratio, weighting pros and cons) and “in-between strategies” (intuition and trust) to come to a decision. The used risk mitigating strategies are present at a technical level and reach from an organizational to an individual level. There are different practices of safeguarding in place such as making the process leading to the decision transparent through documentation (Bio01). [31]

Algorithmic knowledge-based systems are usually applied so that experts can safeguard themselves against “allegations of negligence” (Zinn, 2016: 360) as they can claim to offer “better scientific analysis” through the access to all relevant data (Zinn, 2008: 5). This is a helpful practice for safeguarding a decision within the scientific world as well as within medicine. However, from a legal viewpoint the human experts, usually virologists, still hold the responsibility for the suggestion. This is also reflected within documentation. The report produced by algorithmic tools indicates in a disclaimer that the outcome of the algorithmic tool cannot be trusted alone and must be signed by a human expert. This has among other things to do with the missing approval of the tools as medical device. The validity of the suggestion is then controlled in in vivo observations when physicians see the patient during follow-up control visits and check if the applied therapy works. The virologists get this data about the functioning of the suggested therapy as feedback from the hospitals. If a suggestion within the ML-based tool was given repeatedly, this increased the trust in this very suggestion. The new information produced by the tool must be legitimated as knowledge through the expert panel.

Regarding the different perspectives on risk and uncertainty most of these practices can be regarded from the virologists’ point of view as a “realistic perspective” on risk (Zinn, 2008, p. 4). With this conceptualization in mind the practices above can be seen as “models and scenarios [which] are developed in order to find orientation on how to act rationally regarding an uncertain future even when the knowledge is limited.“ (Zinn, 2008, p. 5)[32] From an organizational perspective the scientific and medical field seem to be well equipped to work side by side on this topic or differently put: The uncertainty mitigating strategies of the scientific world are accepted as sound enough

[31] Although it is known that documentation, e.g. given to health insurances, can have different functions, including misrepresenting what was done in clinical practice to act in favor of the patient (for the difference between practice and documentation see Schubert, 2022).

[32] The limitation of knowledge is here rather the general limitation of knowledge in science and medicine, so to say a knowing that probably not everything is known yet, and not that there is little knowledge available.

for a suggestion to be taken seriously within the medical world. This can also be seen by the fact that virologists are regarded as the experts for the virus and physicians consider their suggestions for their decision about a suitable antiretroviral therapy.

The construction of (un)certainty strengthens virologists' relevance

While we just looked at risk and (un)certainty from the realist perspective of the field, risk and (un)certainty can also be conceptualized as socially constructed. The following section shows what Atkinson already proclaimed in the 1980s. Sometimes the construction of certainty can be more relevant in the field than uncertainty. Nowadays, in 90% of the cases the treatment of HIV constitutes a safe and effective choice (cf. Bio01). Thus, within this case study, the explicit construction of certainty in the field is more prevalent than the construction of uncertainty, risk and danger. In most cases when a common procedure can be followed and an established routine is present, certainty is constructed in the virologists' accounts of the process of successfully figuring out a treatment suggestion, for which their expertise is constructed as necessary. In contrast, when speaking about rare, complicated patient cases which present risky health situations uncertainty is constructed by virologist Bio01.

> *"And then you always try, in dialogue with the doctor and the virologist, the two then always try to approach each other. Then the doctor says, no, but I can't give that one, because he has kidney problems. And the virologist then says, yes, but we have to have a look, maybe we can give this one. No, because the virus looks as if it would tackle it here. We could suppress that with the drug, maybe we try that, and then you approach each other like that. But sometimes we all just fail. And then there are resistant viruses and we can't do anything. That also happens. But that has become rare in the meantime, has to be said."*

These situations can only be overcome when different experts (virologists and physicians) work together and rarely they all fail. Thus, virologists construct themselves as experts through constructing the certainty they have about the treatment of the illness at the same time as they construct uncertainty in risky health situations and show there is a procedure to deal with both. Also, through knowing what represents a risky and what represents a safe situation they construct themselves as experts. Through the construction of different shades of (un)certainty they simultaneously construct their profession as indispensable within the process of finding a treatment suggestion, even if algorithmic DSS are present already since the early 2000s. The uncertainty of the tools still must be transformed into certainty by the virologist. The simultaneous construction of certainty and uncertainty makes sense if both are conceptualized the way Atkinson (1984, p. 954) does based on Schütz' phenomenology as "two modes of attitudes toward knowledge and action" that get constructed by scientists and medical professionals in different situations. There might be uncertainties within scientific knowledge (Fox, 1957) and as Light (1979) would add there can also be uncertainties regarding diagnosis, treatment and response of the patient to the measures. These can be answered with certainty within professional judgement or dogmatism and the latter three uncertainties can be managed with mastering the relevant knowledge (Atkinson, 1984). My experts construct certainty when it comes to their decision in usual cases. Then informational and intrinsic uncertainty can successfully be managed because they assume they and their profession knows everything that is needed to take the right decision that contributes to a safe and efficient treatment suggestion and thus efficient management of danger resulting in little risk. However, to reach this certainty, risk mitigating strategies on individual, organizational and technical level were put into place. Procedures must be followed, such as controlled lab data is used as input, a suggestion of treatment is retrieved from HIV TOS and this outcome is controlled for plausibility. Also, the organization reaches more certainty as these procedures including the relating formalities are followed.

Discussion: HIV TOS as case of precision medicine under the dominant paradigm of evidence-based medicine

Zinn and Olofsson (2019, p. 11) claim that "sociological theories tended to neglect the social middle or meso- level where institutional contexts and individuals combine in new and creative ways" and that this "meso approach to the risk and/or uncertainty practices of everyday life would be in the center of such a diagram where the reality of risk meets the social construction and where the individual and the collective is reproduced through the everyday practices." My analysis tended to this desideratum in analyzing individual, organizational and technical practices of risk and uncertainty mitigation within a case study of precision medicine for HIV positive people. Danger, risk, and uncertainty are framed as real within the field but can be interpreted as socially constructed as they also serve a purpose in the field. The calculated probability of a person to obtain resistant HI viruses can be regarded as real risk, as can the danger of getting the wrong medication and dying because healthcare experts did not suggest the right medication be. However, most of HIV positive people can be treated well with available single tablet regimes of ART. This means the probability of this real risk is also socially constructed, as it is brought in as an argument by virologists and bioinformaticians to underline the necessity of DSS and the relevance of their expertise for the decision of a treatment suggestion. At the same time, as certainty can be constructed when treatment selection follows standard procedure (in most cases), uncertainty is constructed in the few complex cases. The virologists construct themselves as experts who know how to distinguish between these cases and know how to handle them, taking necessary measures in risky situations where decisions can best be reached in communication with physicians.

The virologists in my case study try to manage the risk they take in suggesting a treatment for patients with the tools they created. They have developed tables and rules-based tools to collect, organize and make available their scientific knowledge. They have also developed ML-based tools to find new knowledge to be on an even safer side when deciding about a treatment suggestion. However, the introduction of algorithmic tools in their decision process also means they must keep the tools up to date, safeguard and control them, which in the end leads to additional work of risk mitigation. At the same time, as virologists try to distribute their risk to the algorithmic tools, they must take on and manage the risk that stems from the usage of the tools. Fox (1980) already anticipated in 1980 that technology might lead to more uncertainty in medicine as the interpretation of the outcomes again feeds into uncertainty. Some suggest this resulted in managing uncertainty continuously in EBM clinical judgement (Reed et al., 2016; Timmermans & Angell, 2001). The need for managing uncertainties is reinforced in a field where PM demands to offer the best available targeted treatment based on mechanistic knowledge and when EBM demands that all decisions ideally be based not only on scientific knowledge but on clinical trial (CT)-tested measures.

HIV TOS are an example of PM. However, they were developed in a field where also EBM has been present as dominant paradigm. EBM deals with uncertainty in a "population-centric approach", i.e. it is conceptualizing patients as homogenous and assuming that an average patient response is a sufficiently good predictor of an individual reaction to treatment (McCoy, 2020, p. 28). This mean EBM aims at solving intrinsic uncertainty (what cannot be known) by looking at the complexity of the body whose reactions should be studied in CTs. PM tries to solve informational uncertainty (the problem of unprecise knowledge) by collecting more mechanistic knowledge about individual patients (McCoy, 2020). Both PM and EBM are convinced to offer a framework for optimized patient care within medical world's uncertainty (McCoy, 2020). It is contested if the two rationales of EBM and PM are consolable. Experts who claim their logic is very different would, e.g., mention the difficulty to organize RTCs which would be necessary for EBM based on personalized treatments because of the many different clinical trials one would have to conduct in the absence of a standard treatment.[33] However, many experts agree, that even if the two paradigms work with different rationales, EBM can also include evidence from PM (Nardini et al., 2012).

If we look at the HIV TOS case study, we can see that it is for one thing firmly grounded in the logic of PM: More and individual knowledge seems to be important to manage uncertainty. However, the experts also talk about CTs as way to validate algorithmic tools. In the publications about the HIV TOS, CTs (prospective and retrospective ones) are cited to corroborate the relevance and efficacy of the tools (Liu & Shafer, 2006). Thus, the logic of EBM which is present in medicine seems to be present also in this area of PM. However, the developers of HIV TOS whom I interviewed frame it as an unreachable standard, which results in a paradox that cannot be solved. My argument is that the unsolved uncertainty that arises from the fact that the tools have not been validated within CTs and that none of them have obtained approval adds to the uncertainty the experts feel when using the tool, which needs to be addressed by even more refined practices of mitigation to the related uncertainties. Finally, algorithmic tools, whose initial purpose was to replace humans by offering suitable suggestions, now in turn must be controlled by human experts. Hence, the outcome of the ML-based HIV TOS, which is assumed to be a prediction of the future, remains merely a suggestion with uncertainty attached. The uncertainty is transferred to the expert and results in additional uncertainty mitigation work.

Lenz (2021) similarly shows that while physicians were using digital health tools to reduce uncertainty, they at the same time had to manage new uncertainties that came from the tools which they considered as support. Jacobs et al. (2021) demonstrate that physicians would like to see data from RCTs to trust an AI-based system. Whereas Winter and Carusi (2022) observe that trust in AI also increases when physicians and computer scientists develop a system together. The paradox that RCT results can hardly be generated for tools within PM but are still warranted remains. This situation cannot be solved but the resulting risk is transferred to the experts which manage it with different practices such as plausibility control, documentation and taking on the

[33] Nardini et al. (2012) argue that more importantly some experts within PM would even claim it is not necessary to have phase III clinical trials because the mechanistic information should be sufficient to gain knowledge and certainty.

ultimate responsibility of the decision. To reach certainty within treatment decisions, different types of knowledge are applied. Professional practice means having scientific knowledge, as well as relying on experiential knowledge including intuition when taking decisions. Scamell and Stewart (2014) describe in their study how healthcare professionals stick to formalized guidelines but also use experiential knowledge if opportune and tend to conceal the usage of the latter if necessary. Timmermanns and Angell (2001) showed that especially within EBM an new “research-based uncertainty” is present that comes from considering information technologies and epidemiological data for decisions.

Conclusion

The current study indicates the importance of conducting case study analysis to understand which knowledge decision support systems in medicine and health are based on, which function they have in the field, how they are embedded in and influence their surroundings. The historical analysis helps to follow and understand digitization. Only through the analysis of multiple tools in the field, their interaction and functions could be understood. In the end, the tools are used if they prove to be useful for one of the social worlds and if the benefits outweigh the risks, even if additional work must be done to mitigate risks coming from these tools.

3.4 Künstliche Intelligenz in der Medizin? Intersektionale queerfeministische Kritik und Orientierung[34]

Zusammenfassung

Algorithmen werden als zentrale Akteure der digitalen Transformation gehandelt. Künstliche Intelligenz (KI) wird als Lösung für dringende aktuelle und zukünftige Probleme in der Medizin gerahmt. Der Beitrag geht der Frage nach, wie – oft unbewusst – faktisch diskriminierende Werte sozialer Ordnung in Algorithmen eingeschrieben werden und der weithin beklagte Gender Bias sowie rassistische Diskriminierung fortgeschrieben oder sogar verstärkt wird. Es wird erörtert, wie eine mit KI verbundene Automatisierung von Diskriminierung Ansprüche an ein gleichberechtigtes Zusammenleben vielfältiger und widersprüchlicher menschlicher Existenz erneut breit diskutierbar macht. Im Beitrag werden diese Fragen anhand des Einsatzes von KI bei der Hautkrebs- und der Brustkrebsdiagnose erörtert. Diese werden mit theoretischen und methodischen Zugängen aus der Genderforschung, die sozialen Konstruktivismus, Poststrukturalismus und New Materialism mit Ansätzen der Intersektionalitätsforschung und der Queer Theory verbinden, konfrontiert.

Algorithmen werden als zentrale, maschinelle Akteure der digitalen Transformation gehandelt. Sie werden einerseits für eine Menschen nicht mögliche Neutralität und Objektivität gepriesen. Andererseits aber werden sie auch als potente Instrumente einer gesellschaftlichen Ordnung verurteilt, die aufgrund von Sexismus, Rassismus und Klassismus viele Menschen diskriminiert (Benjamin, 2019; Noble, 2018; O'Neil, 2017). Dieser Beitrag argumentiert, dass die Genderforschung mit der Analyse intersektionaler Prozesse der Privilegierung und Diskriminierung hier Orientierung bietet. Können Algorithmen auch genutzt werden, um Ungerechtigkeit ans Licht zu bringen? Der Beitrag zeigt, wie – oft unbewusst – faktisch diskriminierende Werte sozialer Ordnung in Algorithmen eingeschrieben werden und der weithin beklagte *Gender Bias* sowie rassistische Diskriminierung fortgeschrieben oder sogar verstärkt werden. Es

[34] The publication has been published as follows: R. Baumgartner and W. Ernst. Automatisierte Gerechtigkeit? Kritik und Orientierung für die digitale Transformation. GENDER Zeitschrift für Geschlecht, Kultur und Gesellschaft. 2023, 1, 12-25.

wird erörtert, wie eine solche Automatisierung von Diskriminierung Empörung erzeugt und damit offensichtlich Ansprüche an ein gleichberechtigtes Zusammenleben im Sinne der Gleichwertigkeit aller Menschen erneut breit diskutierbar macht (Alhutter et al., 2020). Im Beitrag wird anhand von zwei Beispielen aus der Medizintechnik (Einsatz von KI bei der Hautkrebsdiagnose und der Brustkrebsdiagnose) gezeigt, wie der Einsatz von KI – entgegen der großen Versprechungen – bisher eher zu einer Verstärkung von Diskriminierung beiträgt. Es wird gezeigt, wie die Orientierung an den theoretischen und methodischen Ansätzen der Gender Studies Hinweise gibt, dieses soziotechnische Problem zu überwinden.

Methodische Grundlagen

Geschlecht wird in der aktuellen Genderforschung als ambivalentes, historisch vielfältiges sowie vielfältig mit anderen Kategorisierungen verknüpftes und widersprüchliches Phänomen des Werdens erforscht (Ernst, 2021). Eine Ausrichtung an antirassistischer, queerfeministischer Geschlechterforschung unter Berücksichtigung von Intersektionalität wird mehr und mehr zum methodischen Standard der Gender Studies. Dabei wird Vergeschlechtlichung als Prozess betrachtet, in dem Geschlecht erst im sozio-kulturellen Zusammenhang hergestellt und ausgedrückt wird. Geschlecht wird als „doing gender“ und „doing difference“ (Fenstermaker & West, 2001) verstanden. Es bezeichnet nicht bestimmte bzw. bestimmbare Eigenschaften, sondern vielmehr eine performative Auseinandersetzung mit vergangenen, gegenwärtigen und vorgestellten zukünftigen Geschlechterordnungen (Butler, 2004). Diese Ansätze ermöglichen das Überwinden eines statisch gedachten binären oder dichotomen Geschlechterbegriffs. Von der prozesshaften Dimension der Vergeschlechtlichung sind Körper nicht ausgeschlossen. Sie werden als Teil einer sozialen, psychischen, intellektuellen und materiellen Auseinandersetzung im unablässigen Gelebtwerden, Geordnetwerden und Regiertwerden erforscht. Auf der Grundlage der Unbestimmtheit von Materie, Zeit und Energie wird auf eine prinzipielle Unbestimmtheit von Identität geschlossen (Barad, 2015). Diese begründet eine immerwährende Erneuerbarkeit der Vorstellungsräume und Erfahrung von Geschlecht. Die Veränderbarkeit, Verbundenheit und Spezifität jeder Geschlechtererfahrung kann so normal und natürlich werden, gerade weil Normalität und Natürlichkeit als Zuschreibungsprozesse entlarvt werden (Ernst, 2021).

Mit dem Begriff „Intersektionalität“ (Crenshaw, 1989) wird Geschlechterdiskriminierung kontextualisiert und tradierte Kategorisierungen des Normalen dekolonisiert. Für diese Auseinandersetzung ist zu beachten, dass „intersectionality is a knowledge project of resistance that aims to bring about change“ (Collins, 2019, p. 289). Wurde Intersektionalität bisher oft missverständlich als additives Set von individuellen, fixierbaren, diskriminierten Merkmalen dargelegt, so ist inzwischen klar, dass Prozesse von Merkmalszuschreibungen in gesellschaftliche Herrschaftsprozesse methodisch relevant gemacht werden müssen: „Ohne den Bezug zu den Herrschaftsverhältnissen und ihren Strukturkategorien, in deren Zusammenhang sie stehen, können die individuellen Erfahrungen lediglich als subjektive Befindlichkeiten und partikulare *handicaps* erfasst werden“ (Klinger, 2013, p. 59; Hervorh. im Original). Methodisch präzisierend wird betont:

> „Prozesse der Feminisierung oder Maskulinisierung sind demnach einerseits in Zusammenhang mit historischen Ereignissen wie denen des Kolonialismus, Imperialismus und des modernen/kolonialen Weltsystems in Beziehung zu setzen und andererseits mit gesellschaftlichen Verhältnissen und institutionalisierten kulturellen Praktiken, in denen hegemoniale Verständnisse in den Alltagsverstand übertragen und von den Subjekten performativ angeeignet und verkörpert werden.“ (Gutiérrez Rodríguez, 2011, p. 89)

Dieses Analysieren vielfältig verwobener Privilegierungs- und Diskriminierungsprozesse zielt auf die strukturelle Überwindung von Diskriminierung. Genderforschung verbindet so sozialen Konstruktivismus, Poststrukturalismus und New Materialism mit Ansätzen der Intersektionalitätsforschung und der Queer Theory. Diese theoretische und methodische Orientierung dient im Folgenden der Analyse des Einsatzes von Algorithmen in der Medizin.

Auf diese Weise kann vielfältige geschlechtliche Existenz in ihrer wandelbaren Verflochtenheit und Uneindeutigkeit für die KI-Forschung geltend gemacht werden. Mit diesem Ansatz können Prozesse der Diskriminierung durch KI entschärft werden. Antirassistische, queerfeministische KI-Forschung kann so angestoßen werden, um Barrieren in und durch KI-Produkte zu identifizieren und wegzuräumen.

Grundlagen und Funktionsweise von KI scheinen diesem Unternehmen bisher entgegenzustehen. KI basiert meist auf sog. maschinellem Lernen (ML), dem Management großer Datenmengen („Big Data“) sowie der Automatisierung von Testmethoden (Prietl, 2019). Isolierung, Quantifizierung und Mechanisierung stellen die basalen methodischen Schritte dar. Daten werden maschinell, automatisiert ausgewertet, in Verknüpfung mit Methoden der Mustererkennung, der Statistik, der

Wahrscheinlichkeitsrechnung sowie einer konditionalen Zukunftsidee. Daher muss analysiert werden, welche Daten überhaupt verfügbar sind, welche Prozesse zuerst isoliert werden müssen, um gezählt werden zu können, und welche Daten relevant erscheinen. Ebenso muss untersucht werden, wer die Muster und Mechanismen vorgibt, nach denen ausgewertet wird, und wie Kriterien bestimmt werden, nach denen sortiert, bewertet, eingeteilt wird (Benjamin, 2019). Dabei werden Probleme offensichtlich: Korrelationen gerinnen zu vermeintlichen Kausalitäten (Noble, 2018); widersprüchliches menschliches Sein und Werden wird zu linearen Vorausberechnungen (Alhutter et al., 2020); veränderliche Unterscheidungen gerinnen zu statischen Differenzen (O'Neil, 2017).

Mit dem oben eingeführten Analyseansatz der Gender Studies, so unser Argument, lässt sich Diskriminierung intersektional präzise aufdecken und die Vermeidung einer Automatisierung von Diskriminierung wissenschaftlich begründen: Es muss bei jedem Datensatz untersucht werden, ob Daten von allen Menschen als gleich wichtig erachtet werden. In der KI-Forschung müssen daher Methoden entwickelt werden, mit denen die lernenden Maschinen von der Vielfalt und Ambivalenz geschlechtlicher Körper und geschlechtlicher bzw. sexueller Lebensweisen unterrichtet werden. Benjamin (2019) macht deutlich, dass Funktionsweise und Einsatz von Algorithmen bislang an einem Modell hegemonialer Männlichkeit orientiert sind. Das heißt, diejenigen, die ohnehin über die meisten Privilegien verfügen, werden erneut privilegiert, während die Diskriminierung jener Personen, die bis heute diskriminiert werden, mit der Automatisierung verschärft und verdeckt wird. Es ist daher notwendig zu untersuchen, was oder wer als Standard, als Norm und Normalität den Rechenautomaten vermittelt wird. Auf die Medizin bezogen ist zu untersuchen, was es den betroffenen Personen nützt, wenn ihr Erkrankungsrisiko von statistischen Mittelwert-Rechnungen abgeleitet wird. Es ist zu fragen, wie sich das Vorantreiben einer radikalen Quantifizierung gesundheitlicher Phänomene und deren Verknüpfung mit automatisierten Prognosen auf die staatliche Verantwortung für gesundheitliche Versorgung auswirkt (Klinger, 2022).

Medizin

Digitalisierung ist schon seit den Expertensystemen in den 1980er-Jahren Thema in der Medizin. Spätestens seit dem Trend der personalisierten Medizin spinnen sich verstärkt Netzwerke mit Informatiker*innen und Bioinformatiker*innen, welche für medizinische Anwendungen, basierend auf neueren Techniken der KI wie ML und Deep

Learning (DL), genutzt werden können. Seitdem tauchen auch in der Medizin Anwendungen auf, in denen diskriminierendes Vorgehen belegt werden konnte. Negative Konsequenzen des Racial Bias und Gender Data Gap werden auch für die Gesundheit diskutiert (Criado-Perez, 2020; Sjoding et al., 2020).

KI wird als Lösung für dringende aktuelle und zukünftige Probleme in der Medizin verhandelt. Dazu gehören scheinbar unlösbare Fragen in der Therapierung von Krankheiten, Alterung der Gesellschaft, Mangel an medizinischem Personal und insgesamt steigende Gesundheits- und Forschungsausgaben, die bisher nicht zu einer besseren allgemeinen Gesundheit führen, sowie gesundheitliche Ungleichheit (Baumgartner, 2021a).

Kritische Stimmen äußern eine Vielzahl an Bedenken gegen den breiten Einsatz von KI für Zwecke der Therapiefindung, Diagnose, Prognose etc. (acatech et al., 2021; Baumgartner, 2021a; Figueroa et al., 2021; Schneider, 2021). Der Technikradar 2021 (acatech et al., 2021) analysiert, welche Aspekte der Gesundheitsversorgung, z. B. das Arzt-Patienten-Verhältnis, sich durch die digitale Transformation wie verändern könnten, und sieht u. a. die digitale Gesundheitskompetenz aller als hochrelevant für eine positive Entwicklung. Figueroa et al. (2021) besprechen, wie wichtig ein intersektional feministischer Blick ist, damit digitale Gesundheit nicht Frauen* benachteiligt. Wenn eine Technik so stark auf Daten basiert, die innerhalb einer gesellschaftlichen Ordnung, die viele diskriminiert und einige wenige privilegiert, produziert wird und die zugehörige Technik von Personen innerhalb dieser sozialen Ordnung entwickelt und validiert wird, ist es unumgänglich, dass sich faktisch diskriminierende und stereotypisierende Werte dieser sozialen Ordnung in Daten und Technik wiederfinden. Wie kann also genau diese Technik am Ende zu gleichberechtigten Ergebnissen führen und fairer sein als das vermeintlich subjektiv entscheidende Gesundheitspersonal? Der Eingang von Stereotypen, Diskriminierung und Ungleichheiten kann auch durch die KI geschehen, die ebenfalls von Menschen mit unterschiedlichen Agenden entwickelt wurden. Dies kann zu Ergebnissen führen, die ähnlich diskriminierend sind wie die derzeitigen auf Menschen gestützten Praktiken, mit dem zusätzlichen Nachteil, dass die Arbeitsweise von KI wenn überhaupt nur für Expert*innen durchdringbar ist. Dazu kommen Herausforderungen, die Technik generell mit sich bringt. Es ist aufwendig, in Technik geronnene Werte und Normen zu analysieren. Digitale Technik kann sehr überzeugend wirken und ein übermäßiges Vertrauen in datenbasierte Technik wird auch vom Gesundheitspersonal befürchtet (boyd & Crawford, 2012; Cabitza et al., 2017).

KI innerhalb der Medizin wird jedoch auch die Möglichkeit zugeschrieben, Diskriminierung aufzudecken und zu minimieren. Vor allem statistikbasierten Systemen wie ML wird das Potenzial zugeschrieben, gesundheitliche Ungleichheiten aufdecken zu können und so zu einer gleichberechtigten Gesundheit zu führen (Topol, 2019b). Dieses Aufdecken gesundheitlicher Ungleichheiten wird in manchen Ethik-nahen Diskursen als ernstzunehmender Vorteil des KI-Einsatzes mit dem geringsten Risiko bewertet (Wachter et al., 2021).

Zurzeit gibt es viele kritische Stimmen, die zwar allgemein auf Herausforderungen, Probleme und Risiken hinweisen, jedoch wenige Studien, die sich mit konkreten Fallbeispielen befassen. Diese sind jedoch von großer Bedeutung, da genau der Einzelfall erst eine Bewertung von Effekten und Risiken möglich macht. Obermeyer et al. (2019) fanden heraus, dass die Verwendung eines falschen Proxys für den Gesundheitszustand zu rassistischer Diskriminierung führen kann. Sie zeigen, wie niedrigere Gesundheitsausgaben für Persons of Color (PoC) in den USA fälschlicherweise darauf zurückgeführt wurden, dass sie gesünder seien, und nicht darauf, dass PoC weniger Leistungen von Gesundheitseinrichtungen erhalten. Studien zeigen, dass auch KI-basierte Anwendungen zur Messung des Blutsauerstoffs oder zur Hautkrebserkennung einen Racial Bias aufweisen (Sjoding et al., 2020; Wen et al., 2022). Criado-Perez (2020) und Cirillo et al. (2020) benennen Gesundheitsrisiken durch fehlende Daten über Frauen (z. B. den Gender Data Gap), welche mit erstarkender digitaler Medizin relevanter werden. Der Fokus auf Daten und Datenlücken ergibt sich daraus, dass in den gängigen KI-Methoden nur Personengruppen, die in Datensätzen vertreten sind, in die Trainingsdaten für die Entwicklung KI-basierter Anwendungen einbezogen werden können. Dies stellt für marginalisierte Gruppen ein Risiko dar, weil sie oft bei Datenaufnahme und Technikentwicklung nicht berücksichtigt werden. Auch für die Aufdeckung gesundheitlicher Ungleichheiten mittels KI müssen Daten entlang von gängigen sozialen Kategorien markiert werden, damit diese retrospektiv entlang dieser Kategorien analysiert werden können. Das heißt, Diversität müsste in den Daten benannt und abgebildet werden. Nur dann ist es möglich, Ungleichheiten zu erkennen, zu benennen und zu ändern. Jegliche Berücksichtigung von Diversität erfordert und fördert daher üblicherweise erstens ein verstärktes Sammeln von Daten und zweitens das Zuweisen von gängigen sozialen Kategorien zu eben diesen Daten (Baumgartner, 2021a, 2021b). Mehr Daten zu marginalisierten Gruppen zu sammeln kann, wenn gängige Kategorien zur Klassifizierung verwendet werden, zu einer Reifizierung dieser Kategorien führen (Baumgartner, 2021a). Unabhängig davon sind Standardisierung, Quantifizierung und Kategorisierung wirkmächtige Prozesse innerhalb der Medizin, mit denen die Medizin schon lange zur Naturalisierung von Kategorien beiträgt. In Kombination wird eine datafizierte und digitale Medizin daher zu einer Verfestigung von Kategorien führen, so die Befürchtung von Expert*innen (Baumgartner, 2021b; boyd & Crawford, 2012).

Im Folgenden analysieren wir zwei Beispiele aus der Onkologie, einem Anwendungsbereich der Medizin, in dem große Hoffnungen auf KI gesetzt werden. Beide Indikationsgebiete bieten zudem spannende Fallbeispiele für eine intersektionale und queerfeministische Analyse.

Brustkrebs

Der Einsatz von KI zur Früherkennung von Brustkrebs wird hier im Zusammenhang mit verschiedenen Narrativen zu Brustkrebs betrachtet (Nielsen, 2019). Dabei ist zu fragen, wie die Brust von einem macht- und lustvollen Körperaspekt in verschiedenen Narrativen zu einem sich der eigenen Kontrolle entzogenen, angstbesetzten, verdächtigen, todbringenden Aspekt umgewertet wird. Audre Lorde zeigt in ihrem berühmten *Krebstagebuch*, wie sie trotz der eigenen Diskriminierung als lesbische Schwarze Frau im profitorientierten US-amerikanischen Gesundheitssystem ermächtigende Aspekte in sich selbst findet, um informierte, medizinisch relevante Entscheidungen zu treffen (Lorde, 2000). Lorde kritisierte schon 1979 die mangelnde Brustkrebsvorsorge, die mangelnde Erforschung „umweltbedingter“ Ursachen, die Orientierung der Amerikanischen Krebsgesellschaft an der Mastektomie als einziger Behandlungsmethode und den faktischen Zwang zum anschließenden Tragen einer Brustprothese oder eines Silikon-Gel- Implantats (Lorde, 2000, p. 60ff). Während ihres 14-jährigen Überlebens erarbeitete sie eine intersektionale queerfeministische Orientierung für die Krebsforschung, bei der Selbstermächtigung und Solidarität zentral sind. Diese Orientierung stellt den Maßstab für die folgende Analyse von KI bei der Brustkrebsdiagnose dar.

Vor zwanzig Jahren stellte das Nachrichtenmagazin *Der Spiegel* Brustkrebs als „Katastrophe für die Frauen“ (Stockinger, 2002) dar: „Nach Schätzungen des Bonner Gesundheitsministeriums sterben in Deutschland jährlich rund 4 000 Frauen, weil bei ihnen der Knoten in der Brust nicht rechtzeitig diagnostiziert worden ist“ (Stockinger, 2002, p. 204). Medizingeschichtlich betrachtet steht es um die Brustkrebsheilung in Deutschland im Jahr 2018 deutlich besser: „Gut 70 Prozent der Brustkrebspatientinnen werden brusterhaltend operiert und erhalten anschließend eine Strahlentherapie. Wird Brustkrebs frühzeitig erkannt, gilt er zu 80 Prozent heilbar“ (Hombach, 2018, o.S.). Inzwischen hat sich die Mammografie als Standard in vielen europäischen Staaten etabliert: Entsprechende Krebsfrüherkennungsprogramme gibt es in Deutschland seit 2009. Dabei ist die Mammografie umstritten (Helsana, 2017). Neben längeren Wartezeiten auf den Befund durch Radiolog*innen sind falsche positive Befunde und falsche negative Befunde häufig. Die Empfindung der Mammografie als unangenehm bis schmerzhaft sowie die Strahlenbelastung werden von vielen Frauen* beklagt. Daher ist die Teilnahme freiwillig (BM für Gesundheit, 2016).

Im Januar 2020 pries nun eine Publikation des renommierten US-basierten Wissenschaftsmagazins *Nature* die Fortschritte der KI zur Interpretation von Mammografien: „Here we present an artificial intelligence (AI) system that is capable of surpassing human experts in breast cancer prediction“ (McKinney et al., 2020, p. 89). Bei genauem Lesen zeigt sich allerdings, dass zwar eine Verbesserung hinsichtlich der Reduktion der Fehldiagnosen bei der Vorhersage von Brustkrebs im Vergleich zu früheren Versionen von Algorithmen erreicht werden konnte: „We show an absolute reduction of 5.7 % and 1.2 % (USA and UK) in false positives and 9.4 % and 2.7 % in false negatives“ (McKinney et al., 2020, p. 89). Die KI erwies sich aber nicht besser als das Ergebnis, das durch den europäischen Standard der Konsensus-Interpretation von zwei Radiolog*innen erzielt wurde (McKinney et al., 2020, p. 92).

Diese Publikation löste viele euphorische Medienberichte über die Chancen von KI aus. Demgegenüber wies die Deutsche Statistische Gesellschaft unter dem Titel „Unstatistik des Monats Januar“ vom 01.02.2020 auf das Problem hin, „wie AI-Erfolge in der Presse übertrieben werden und die Frage nach dem Nutzen für Patientinnen und Patienten nicht gestellt wird“ (Gigerenzer & Weiler, 2020, o.S.). Weiter wird problematisiert: „[J]e besser die Diagnose-Systeme werden, desto mehr kleine und klinisch irrelevante Krebsformen werden entdeckt, die nur technisch gesehen Krebs sind“ (Gigerenzer & Weiler, 2020, o.S.). Das heißt, eine für die Brustkrebsdiagnose ganz wesentliche Unterscheidungsfähigkeit ist bei KI nicht gegeben, das führt zur sogenannten Überbehandlung. Die Autor*innen kommen auch in Bezug auf die Bewertung von Mammografie zu einem ernüchternden Fazit: „Von je 1 000 Frauen, die nicht zum [Mammografie-] Screening gehen, [sind] nach 10 Jahren etwa 5 an Brustkrebs gestorben; mit Screening sind es 4“ (Gigerenzer & Weiler, 2020, o.S.). Die mediale Überbewertung von Entwicklungen in der KI-Forschung kann zu einer Fehleinschätzung der Potenziale von KI führen und damit zu einer Fehlinvestition von Forschungsmitteln.

Das folgende Beispiel zeigt, wie Frauen*, ihre Sorge um ihre Gesundheit und die Abbildungen ihrer Körper als profitables Forschungsmaterial begehrt werden. Wenige Monate nach der Veröffentlichung in *Nature* gründete sich die HIPPO Foundation mit der Kampagne „Victoria 1.0“. In einer weit verbreiteten Internet-Brustkrebskampagne erfolgte ein eindringlicher Aufruf zur Datenspende (Röntgenbilder einer Mammografie, Magnetresonanztomografie (MRT) oder pathologische Bilder). Das Anliegen wurde als „Open-Source-Ansatz“ bezeichnet: „ein daraus entstehendes KI-Modell offenzulegen und damit der Allgemeinheit zur Verfügung zu stellen“ (Klöckner & Rybicki, 2020, o.S.). Das Ziel der Kampagne wird einerseits als Solidarität mit potenziell von Brustkrebs Betroffenen beschrieben und andererseits als europäisches Wettbewerbsziel: „Er sieht in dem Open Source Ansatz einen Weg für Europa, den Rückstand im Bereich KI im Gesundheitswesen gegenüber den USA und China aufzuholen. ‚Wir erleben, wie dort Firmen mit gigantischen Datensätzen KI-Modelle aufbauen‘, sagt [‚Bart de Witte, Gründer der Hippo Foundation, früherer Digital-Health-Vorstand beim Softwareunternehmen IBM in der DACH-Region‘]“ (Klöckner & Rybicki, 2020, o.S.). Auf der Website der Stiftung wird entsprechend der wirtschaftliche Gewinn durch Start-ups gepriesen.

Obwohl das Problem der Fehlerhaftigkeit bei der Auswertung von Mammografien trotz der Vorgabe staatlich garantierter Qualitätssicherung von zwei ausgebildeten Personen bekannt ist (BM für Gesundheit, 2016), wird nichttechnologischen Alternativen kaum Beachtung geschenkt. Seit 1983 werden blinde Frauen aufgrund ihres überdurchschnittlich guten Tastsinns zu „Brustuntersuchungsschwestern“ (Hengstenberger, 2005, p. 46f) ausgebildet. Eine Studie der Universität Erlangen belegt die Akkuratheit ihrer Diagnose als vergleichbar mit entsprechend ausgebildeten Ärzt*innen (Lux et al., 2019). Die untersuchten Teilnehmerinnen* (98 %) wollten die Untersuchung durch die medizinischen Tastuntersucherinnen* (MTUs) weiterempfehlen. Die Einstellung von MTUs könnte also die Beteiligungsquote an Brustkrebsfrüherkennungsprogrammen deutlich erhöhen: „Including MTEs could lead to benefits in healthcare and breast diagnosis, while also generating occupational opportunities for visually impaired people“ (Lux et al., 2019, p. 45). Auf diese Weise, so wird

argumentiert, kann eine Behinderung zur Begabung werden und auf der Basis solider beruflicher Anstellung die Gesundheit aller Frauen fördern – im Sinne von Selbstermächtigung und Solidarität.
Zusammenfassend muss die Euphorie um die Entwicklung von Brustkrebsdiagnosealgorithmen kritisch betrachtet werden. KI-gestützte Brustkrebsdiagnose, so geht aus dem Vorangegangenen hervor, scheint für Frauen* und potenzielle Patient*innen kaum nützlich. KI in der Brustkrebsfrüherkennung kann bislang nicht als Entlastung für Radiolog*innen betrachtet werden. Weiterhin erscheint die Motivation zur Datenspende und Mitgliedschaft in einem Datenpool aufgrund einer Internetkampagne einer privaten Organisation eine fragwürdige Art der Partizipation, Solidarität oder Ermächtigung. Teilhabe scheint hier auf die Ablieferung von Daten- und Bildmaterial reduziert und keine Mitbestimmung über die Verwendung des Materials oder des wirtschaftlichen Gewinns zu beinhalten. Dagegen erscheint die Ausbildung von blinden Frauen* zu medizinischen Tastuntersucherinnen aufgrund der großen Akzeptanz und des Diagnoseerfolgs vielversprechend.
Mit dieser Kontextualisierung des Einsatzes von KI zur Brustkrebsdiagnose werden verschiedene Positionierungen zur Krankheit deutlich. Es muss dringend untersucht werden, was genau von der KI auf der Mammografie als potenzielles Karzinom im Brustgewebe identifiziert wird und inwiefern die Tumorerkennung auf eine Rechenleistung reduziert werden kann. Weiterhin muss untersucht werden, wie prekär der Spielraum zwischen Beteiligung und Ablehnung für potenziell Betroffene ist und wer in diesem heiß umstrittenen sozio-technischen Forschungsfeld eigene Präferenzen geltend machen kann. Lordes *Krebstagebuch* bietet eine bis heute relevante Auseinandersetzung mit Brustkrebs im Zusammenhang mit sich verändernden Bildern und Praktiken von antirassistischer, queerfeministischer Weiblichkeit und Solidarität. Der Film *The Cancer Journals Revisited* (2018) von Lana Lin macht dies anschaulich (Queertactics, 2021). Die Brust ist ein Körperaspekt bzw. Organ, das Ambivalenzen der Un_Sichtbarkeit verdeutlicht – als Gendermarker. Brustkrebs ist eine Krankheit, die diese Ambivalenzen sichtbar macht und verhandeln lässt – im Sinne von „queering bodies" und „undoing gender" (Butler, 2004). Diese Ambivalenzen verdeutlichen medizinische bzw. kulturelle Geschlechterpolitik. KI in der Medizinforschung verschiebt die Forschung von Fragen der Prävention und Behandlung zu Fragen der Prognosen und Risikoberechnung. Datenspende ist eine sehr reduzierte Form der Partizipation und kann vor dem Hintergrund von Privatsphäre und Datenschutz der Beteiligten

sowie dem Verdacht auf Kommodifizierung auch kritisch gesehen werden. Partizipation an medizintechnischen Verfahren muss daher neu diskutiert werden, auch im Hinblick auf ein „informed refusal“ (Benjamin, 2016), ein informiertes Ablehnen von Beteiligung. Diese Orientierung soll zu einer umfassenderen Untersuchung der Möglichkeiten und Grenzen von KI bei der Brustkrebsdiagnose inspirieren, ohne dabei mögliche Alternativen aus dem Blick zu verlieren.

Hautkrebs

Die Haut spielt eine bedeutsame Rolle in der Markierung von und Zuweisung zu sozialen Kategorien. Ihre Farbe wurde von Kolonialist*innen im späten 18. und frühen 19. Jahrhundert als Marker für rassifizierende Unterscheidungen gesetzt und gilt nach wie vor als Angriffspunkt für Rassifizierung (Kuria, 2014; Terhart, 2014).
und die Art, wie wir uns um sie kümmern, kann nicht nur als Ausdruck, sondern auch als Marker für Geschlecht, sexuelle Orientierung, Klasse und Alter herangezogen werden. Dies geschieht in unterschiedlichsten intersektionalen Überschneidungen: Make-up von Frauen* und Männern* unterschiedlicher sexueller Orientierung, geschlechtlicher Vielfalt oder sozio-ökonomischer Positionierung, gebräunte oder weiße Haut (Aufhellungscremes) als Marker für soziale Klasse bzw. sozialen Status und Lebensstil etc. In diesem Sinne greifen wir hier poststrukturalistische Geschlechtertheorie, Sozialkonstruktivismus und intersektionale Ansätze auf, um die bisherige Thematisierung von Geschlecht zu problematisieren.
Hautkrebs ist in Deutschland die häufigste Krebserkrankung. Das Auftreten steht im Zusammenhang mit (ungeschützter) UV-Lichtexposition. Nur ein frühes Erkennen kann die erfolgreiche Therapie sicherstellen (RKI - Robert Koch Institut & Zentrum für Krebsregisterdaten, 2021). Damit sind Vorsorgeuntersuchungen, auf die z. B. in Deutschland Krankenversicherte zweijährlich Anspruch haben, höchst relevant. Die Dermatologie misst daher dem individuellen Verhalten große Relevanz bei und diskutiert einen Zusammenhang von Hautkrebs und sozialen Kategorien.
Das RKI (2021) weist darauf hin, dass Männer ein fast zweifach höheres Risiko haben, an einem Melanom zu sterben. Generell erkranken Männer etwas häufiger an Hautkrebs als Frauen. Dafür werden unterschiedlichste Gründe herangezogen: Männer seien eher der Sonne ausgesetzt, verwendeten dabei weniger Sonnenschutz und nähmen weniger oft hautärztliche Kontrollen in Anspruch. Hautkrebs tritt bei Männern eher am Rumpf und bei Frauen eher an den Gliedmaßen auf, was das selbsttätige

Erkennen von Hautveränderungen für Männer erschwert. Frauen würden ihre Haut mehr beobachten als Männer. Die Beschaffenheit der Haut von Männern begünstige überdies eine UV-Licht-schädigende Wirkung. Höhere Östrogenwerte und das aktivere Immunsystem von Frauen hätten zudem einen günstigen Einfluss bei der Hautkrebstherapie (Stallings, 2020). Sexuelle und geschlechtliche Minderheiten sowie Personen mit Östrogenhormontherapie hätten ein erhöhtes Risiko, Hautkrebs zu entwickeln (Marks et al., 2020). Das Beispiel zeigt, wie im Falle von Hautkrebs zwar biologische, aber vor allem auch soziale Faktoren wie geschlechtsstereotypes Verhalten als Hypothesen für die unterschiedliche Auftrittsfrequenz besprochen werden. Wie können Menschen entlang unterschiedlicher sozialer kategorialer Zugehörigkeit vor dem Erkranken und Sterben an Hautkrebs geschützt werden und gleichzeitig ihre diversen Zugehörigkeiten fluide gedacht werden?
Bereits Anfang 2000 wurden computergestützte Instrumente zur Erkennung von Hautkrebs entwickelt. Ziel war u. a., dass Lai*innen ihre Haut selbst kontrollieren können und damit die Dermatologie entlasten. Apps zur Hautkrebserkennung sind inzwischen verfügbar (Vienna online, 2022). Kommerziell erhältliche computergestützte Diagnosetechniken behaupten, genauso gut oder sogar besser zu sein als Dermatolog*innen (del Rosario et al., 2018; Esteva et al., 2017). Anfang 2020 jedoch bekam Racial Bias in KI-Tools zur Hautkrebserkennung vermehrt Aufmerksamkeit in wissenschaftlicher und medialer Berichterstattung. Was war passiert? Esteva et al. (2017) trainierten ihre KI vor allem mit Bildern von Hautläsionen auf hellen Hauttypen. Melanome sehen jedoch auf verschiedenen Hauttypen unterschiedlich aus. Um Muttermale auf unterschiedlicher Haut erkennen zu können, müssten Bilder unterschiedlicher Hautfarben mit unterschiedlichen Arten von Muttermalen in den Trainingsdaten enthalten sein. Das würde allerdings bedeuten, dass diese Bilder vorhanden und korrekt bezeichnet sein müssten. Wen et al. (2022) führten eine systematische Überprüfung von „publicly available skin cancer image datasets“ (Wen et al., 2022, p. e64) durch, die üblicherweise als Trainingsdaten für solche Anwendungen verwendet werden. Sie fanden „a substantial under-representation of darker skin types“ (Wen et al., 2022, p. e64). Nur 2,1 % lieferten Daten zum Fitzpatrick-Hauttyp. Die Fitzpatrick-Skin-Type-Skala (FST) „is the most commonly used classification system in dermatologic practice“ (Ware et al., 2020a, p. 77). Sie wurde 1975 entwickelt „to assess the propensity of the skin to burn during phototherapy“ (Ware et al., 2020a, p. 77). Laut Wen et al. (2022) haben gängige Hautdatensätze eine „limited applicability to real-life clinical settings

and restricted population representation, precluding generalisability. Quality standards for characteristics and metadata reporting for skin image datasets are needed“ (Wen et al., 2022, p. e64). Einige Artikel schlagen technische Lösungen für das Problem von Racial Bias bei KI-gestützter Hautkrebserkennung vor. „ML software algorithm [sic] that is trained to recognize melanoma on all skin types“ wären laut Adamson und Smith (2018, p. 1247) das Ziel. Das (2021) beschreibt diesem Vorschlag folgend eine KI-Architektur, die die Relevanz der Hautfarbe im Modell verstärkt, um sie besser zu berücksichtigen. Das weist jedoch auch auf die Datenlücke bei Bildern mit dunklerer Hautfarbe hin und bezieht mit der Begründung, Menschen mit diesen Hauttönen „are significantly less likely to get melanoma“ (Das, 2021, p. 1720), keine Bilder der dunkleren Fitzpatrick-Skin-Types IV–VI in die Analyse ein. Es ist jedoch bekannt, dass Menschen mit sehr dunkler Haut auch gefährdet sind. Beobachtungsstudien zeigen, dass der Schutz, den sie aufgrund ihres Hauttyps haben, keine mildernde Wirkung auf ihre Sterblichkeitsrate durch Melanome hat (Noor, 2020; Ward-Peterson et al., 2016). Das Hauptproblem in diesem Beispiel scheint die Datenlücke bei Bildern zu dunklen Hauttypen zu sein. Hat dies mit einem Racial Bias bei der Melanomerkennung zu tun, der die Annahme begünstigt, Menschen mit dunkler Haut wären vor Hautkrebs geschützt? Auch die Verwebung der FST mit Rassismus scheint vielfältig: Dunkle Haupttypen (V–VI) wurden erst später zur FST hinzugefügt und die Skala funktioniert für ihren ursprünglichen Zweck der Klassifizierung der Wahrscheinlichkeit von Hautschäden durch UV-Lichtexposition besser bei weißer Haut als bei PoC. Auch kam es bald nach der Entwicklung zur Zweckentfremdung für die Beschreibung verschiedener Hautfarben (und damit auch von *race*). Wie kann Rassismus bei der Hautkrebserkennung verhindert werden, ohne rassistische und Rassismus festschreibende Skalen wie die FST zu verwenden? Es fragt sich auch, wie ein Spektrum aller Hauttypen mit ihrer Fluidität abgebildet werden kann, wer von unterschiedlichen KI-basierten Systemen profitiert und ob sie in heiklen rassifizierenden Bereichen wie diesem überhaupt Anwendung finden sollen.

KI kann in beiden Beispielen (Brustkrebs- und Hautkrebserkennung) als Brennglas für Herausforderungen in der Medizin begriffen werden. Existierende Probleme innerhalb der Medizin bekommen erneut Aufmerksamkeit und sollten im breiten Diskurs besprochen werden. Wofür wollen wir KI einsetzen und mit welchen Mindestanforderungen? Welche Kategorisierungen oder Systeme können die komplexe, sich ständig ändernde Realität am ehesten repräsentieren? Wie viel Datenspeicherung ist notwendig und welche Bereiche und Individuen/Gruppen sollten vor ihr geschützt werden?

Orientierungen für die KI-Forschung

Theoretische Positionen, die im Rahmen von Diversität und Fairness in KI herangezogen werden, basieren auf unterschiedlichen Ansätzen der Gender and Diversity Studies. Einerseits gibt es die Vorstellung, dass nur Kategorien, die verwendet werden, auch sichtbar sind und somit mehr Daten insgesamt und auch Informationen zu Kategorien gesammelt werden sollten. Andererseits gibt es eine Kategorien-kritische Position, die Anstoß nimmt an einer weiteren Naturalisierung von Kategorien. Im Rahmen dieses Beitrags haben wir diese Positionen diskutiert und verortet. Im Verlauf des Beitrags haben wir Möglichkeiten und Denkrichtungen aufgezeigt, wie diese Positionen verknüpft werden könnten, z. B. indem Flexibilität und Veränderbarkeit innerhalb und zwischen den Kategorien mitgedacht werden. Im Beitrag wird deutlich, wie Positionen des New Materialism, der Intersektionalität, der Queer Theory und des Sozialkonstruktivismus für die kritische Analyse algorithmischer Systeme in der Medizin wegweisend sein können.
Betrachten wir diese Analysen vor dem Hintergrund eines intersektionalen Verständnisses von Geschlecht als Phänomene des Werdens (Ernst, 2021), stellen sich zwei umfassende Anforderungen an die KI-Forschung. Erstens: Die Orientierung an der Erkenntnis von Ambivalenz, wie sie in den Gender Studies immer mehr methodisch geltend gemacht wird, kann eine Automatisierung von Stereotypen in der KI-Forschung vermindern. Hierfür erscheint es notwendig, geschlechtliche Vielfalt und menschliche Ambiguität auch in unvorhersehbaren Verbindungen und intersektionalen Zusammenhängen von Privilegierung und Diskriminierung zu verstehen. Eine solche kontextuelle Einbettung entlarvt die reduktionistische Kategorisierung von scheinbar ‚objektiven‘ materiellen Merkmalen als an Stereotypen orientierte Festschreibung von tradierten bzw. längst überwunden geglaubten Hierarchien. Daher sind grundlegende Begriffe und Methoden der KI-Forschung zu überdenken. Es muss geprüft werden, ob in der KI-Forschung Differenzen durch simplifizierende Kategorien erst hergestellt werden und ob Differenzierungen nur statisch, unveränderlich gedacht werden. Stattdessen gilt es, Differenzierungen – realistischer – fluid, vielfältig und überschneidend auszulegen.

Zweitens: Die KI-Forschung muss geschlechtliche Vielfalt und menschliche Ambiguität in unvorhersehbaren Verbindungen und intersektionalen Zusammenhängen von Privilegierung und Diskriminierung methodisch aufgreifen. Partizipatorische Methoden, die in der feministischen Technikforschung entwickelt wurden, könnten auch für die KI-Forschung eingesetzt werden (Suchman, 2019). Suchman schlägt darüber hinaus engagierte transformative Interventionen vor. Solche partizipativen Ansätze und transformativen Interventionen wurden schon im Maschinenbau getestet (Ernst, 2017). Ähnlich fordert Bath (2014) eine Orientierung der Informatik an einem „diffractive design“, bei dem gerade die bisher gesellschaftlich ausgeschlossenen Anderen Orientierung und Standard für innovatives virtuelles Design bieten. Es ist zu prüfen, ob diese Ansätze auf transformative Interventionen in Bereichen übertragbar sind, die von mathematischen Modellen, Algorithmen, digitalen und statistischen Verfahren des Kategorisierens und Zählens strukturiert werden, in deren Funktionsweise nur wenige Einblick haben und die teilweise auch durch Expert*innen nicht mehr durchdringbar sind. Dabei wäre zu klären, ob und ggf. wie KI-Forschung ohne Eindeutigkeiten arbeiten oder Fluidität personaler Identität und Interaktion abbilden kann. Gleichzeitig müsste das Vorgehen der KI in jedem Schritt für Menschen transparent gemacht werden. Dies scheint eine wesentliche Voraussetzung dafür zu sein, realitätstaugliche menschenzentrierte KI zu entwickeln.

Die Zukunft wird zeigen, ob eine in diesem Sinne antirassistische, queerfeministische KI-Forschung denkbar und sozio-technisch umsetzbar ist.

4. Further analysis and comparison of the case studies skin cancer detection tools and HIV TOS

This section will provide a short summary of the conducted studies and the published papers and will offer the theoretical background to relevant categories for the reanalysis and comparison of the two case studies.

HIV-treatment optimization tools were the first case study I researched empirically. I conducted 6 expert interviews with 5 experts from the involved social worlds of virology, bioinformatics, and medicine, most of whom were taking part in the development of AI-based tools (Baumgartner, 2023a, 2023b). Additionally, I analyzed scientific publications about the tools at hand. I coded interviews and publications according to Grounded Theory and used messy maps, as well as situational maps according to Situational Analysis which were adapted for the purposes of understanding what happened in the social arena of HIV TOS development (Baumgartner, 2023a; Clarke et al., 2018; A. Strauss & Corbin, 1994).

Applying a sociology of technology framework, I soon found that a whole family of tools exist. Initial algorithmic tools are based on rule-based algorithms while newer tools are based on machine learning methods very closely related to statistics (Baumgartner, 2023b). Two of the papers in this doctoral thesis report the results from this case study: The paper "Precision medicine for HIV therapy: A tale of successful risk and uncertainty management?" uses a sociology of risk and uncertainty perspective and theories on uncertainty from medical sociology to analyze the different functions the tools have in the field based on the technology employed. It also shows how the users of the tools (virologists) construct uncertainty and certainty to establish their relevance in the field (Baumgartner, 2023b).

In the paper "Personalized HIV Treatment: Bringing marginalized patients to the forefront with situational analysis", I discuss how the social worlds act on distinct but also on shared goals. People with HIV that seek antiretroviral treatment are at the same time centered (being the main reason for the development of the tools) and marginalized (as they are not able to weigh in on their own social construction within this social arena). Different social worlds discursively construct patients in distinct ways: the social world of physicians having direct contact with patients construct them as complex, human people to be treated, the virologists who use the tool and only have contact with blood samples of patients construct them as objects to be examined and treated, while bioinformaticians who only have access to the patients' data construct them as carriers of data (Baumgartner, 2023a).

The paper precision medicine (PM) vs. digital phenotyping (DP) provides an overview of PM such as HIV TOS and digital phenotyping which is a currently developed type of precision medicine for mental health. DP and PM– both show similarities as they apply the same logic of "trust in digitalisable data" (Baumgartner, 2021b, p. 9), even though they rely on different types of data – PM uses validated biological data, whereas for DP, big data from web usage and wearables is used.

The case study on skin cancer detection is based on a scientific literature search applying a feminist science and technology perspective to understand what leads to racial bias in skin cancer detection tools. The field discusses various known categories as relevant to how they lead to higher skin cancer risk. Skin color and race is prominently discussed in these accounts. The final hypothesis of the paper is that the racial bias in the field results in missing data for training AI-based tools in an inclusive way. This again might lead to a perpetuation of racial bias through AI-based tools, which do not work equally well for all skin colors. The case study on breast cancer detection also published in this collaborative paper was analyzed by my colleague Waltraud Ernst and will thus, not receive much attention in this framework (Baumgartner & Ernst, 2023).

One major endeavor of this framework is to compare the two case studies regarding categories used in the AI-based tools and in the field surrounding the tools. During the process of reanalysis, it soon became clear that skin cancer detection tools were the more promising case to be centered, since social categories were made more relevant within the tool. In these tools, the category of race plays a predominant role. Therefore, the following section will further elaborate on the theorization of the category of race in sociology before coming to the reanalysis and comparison of the two case studies.

4.1 The category of race

While feminist STS, which was used for my previous analysis of skin cancer detection tools, has rather essentialist views on race which Brubaker would call "groupism", the sociology of categorization is interested in how categorization and categories work (Brubaker, 2009, p. 28; Müller, 2014). From this perspective race is seen as a contingent observation scheme[35], an epistemology and not an ontology of the world (Brubaker, 2009; Müller, 2014).

Historically, most theorists date the invention of race as a product of colonialism back to the 19th century (Hacking, 2005). Colonizers based the ethics and rationale for colonializing groups of people in foreign countries on the understanding of the inhabitants of these countries not only as different from but also as inferior to them (Müller, 2014). The understandings of difference were backed up by (racist) race sciences during colonialism that were determined to discover differences in traits apart from the obvious visual differences used to distinguish between races (Hacking, 2005; Hirschauer, 2021). While there was from then onwards a common understanding that humans can be divided in different races[36], the way people are categorized in different subcategories varies according to historic phases, geographic regions and social contexts (Hacking, 2005; Müller, 2014).

Institutionalized racial classifications such as early forms of the census were then not only a practice to count people based on different races but aimed to control, "enslave", and "exploit" them (Hacking, 2005, p. 104). The historic example of racial classifications under apartheid in South Africa shows how a racial classification system can become part of a working infrastructure. People were classified based on appearance: skin color (using scales to compare the color of the skin), hair structure, eyes, bone structure, and features. This led to the naturalization of a problematic political category (Bowker & Star, 2000d).

In their function of quantifying groups of people, censuses could also provide more visibility and power to groups they count (Lee, 1993; Scheuerman et al., 2020). Thus, the question of which subcategories are included in the census, including the flexibility in choosing between them, are contested until today (Lee, 1993). Because of their importance and far-reaching effects, racial classifications in censuses have been researched extensively within the sociology of classification and related sociologies (Lee, 1993). Official racial classification systems such as the census are relevant also for other disciplines including medicine and healthcare which rely on these scales when they collect data about subcategories (Epstein, 2007d). The more data-informed

[35] in German "kontingente Beobachtungsschemata" (Müller, 2014, p. 405).

[36] This is no longer accepted in the same way today, based on its refutation in biological and cultural science research (Epstein, 2007d; Roberts & Rollins, 2015).

our chosen perspectives on the world are, the higher the importance of the classification systems used to collect data for further analysis (Heintz, 2021b).
While medicine uses existing classification systems to count people, it also leads to the naturalization of the (sub)categories it counts. Brubaker (2009, p. 32) points out how biomedical research has "important implications for the study of ethnicity, race, and nationalism" because of its "tendency to naturalize social categories". Medicine and healthcare have been a prominent field to research processes of classification, their standardization, and the challenges that accompany them. Epstein (2007b) tended to a multitude of questions around categories and standards in medicine including race in medicine which will be summarized in the following section.

4.2 Race in medicine

Even if science, which in colonialist states involved evolutionary biology, statistical anthropology, and later genetics, tried to find difference within racial subgroups, there is no convincing biological basis for racial classifications to date (Epstein, 2007d; Fenstermaker & West, 2001; Hacking, 2005). This was a challenge for medical research on race which needed scientific facts to use race as a biomedical category. Epstein (2007d) shows how the usage of the category race in medicine was contested starting from the 1990s because of missing scientific evidence. Due to the downsides of the usage of race on the individual and societal levels, Epstein suggests a cautionary use of the category, only in the presence of statistical significance and meaningful explanation (2007d).
Epstein (2007b) also provides a historical analysis of the tendencies to account for social categories in medicine. For a long time, the idea of a 'standard human', a one-size-fits-all approach, was the basis for biomedical research and for the medical treatment of people (Epstein, 2007a). Reformers including feminists, inspired by minority rights movements, started to fight for more precise measurements of biomedical differences between groups of people in the 1980s, to account for diversity and inclusiveness (Epstein, 2007c). Their incentive resulted in the institutionalization of these values in the form of laws and guidelines, newly established federal offices, and review boards at universities and hospitals in the U.S. and found its way into procedures and work practices. Epstein (2007c, p. 6) calls the connected paradigm the "inclusion-and-difference paradigm." These categories are often recorded based on classification systems that are already used by other government agencies, such as the ones used in the

census. However, those have politics behind them and might not even be the most relevant categories to search for biomedical differences (Epstein, 2007a).

Epstein (2007d) discusses several unintended consequences of the inclusion-and-difference paradigm, some being that the differentiation between categories leads to the assumption of real differences between categories and makes it more difficult to generalize knowledge from one category to the other. Moreover, because medicine is interested in biomedical differences, it will ultimately naturalize them, which can then be used in pejorative ways (Epstein, 2007d). He also contends that these naturalized understandings about categories will also spark the interest in and need for other disciplines such as natural sciences and social sciences to tend to the question of inclusion of categories and their biomedical differences (Epstein, 2007c). In the end research with the focus on finding biomedical differences between social categories can change our social understanding of these categories, e.g., framing them as biomedical rather than as socially constructed (Epstein, 2007b).

In a more practical way this paradigm also encourages the development of specific medications for specific sub-categories. Thus, medications targeted for specific races could foster "inaccurate and problematic understandings of race" because they support the wrong assumption that a category is homogenous and disguise "complex sociopolitical processes" needed to assume a certain racial identity (Epstein, 2007d, pp. 227-228). It additionally reinforces the idea of racial differences and might even lead to "dangerously inaccurate understandings of the causes of health disparities": when "racial differences are attributed to biology", other explanations of these disparities might not be found (Epstein, 2007d, p.228).

While Epstein (2007a) acknowledges the need to understand discrimination based on known categories, he is critical about tracking them purely to understand biomedical processes and suggests the study of race in medicine as follows: Instead of simply using known categories to find biomedical differences, we should find the most meaningful categories to research them and identify when it is it alright to assume that there is no meaningful biomedical difference. We should find out how the experience of racial discrimination interacts with biology and the way in which it influences diseases (Epstein, 2007d).

How is this relevant for AI-based tools which work with digitalized data? The following section will take Epstein's work as well as the before mentioned theorization on the sociology of categorization into account for the reanalysis of the two case studies on skin cancer and HIV TOS.

4.3 AI-based skin cancer detection: Skin cancer as consequence of biology or the social?

Skin cancer is a highly prevalent form of cancer (RKI, 2021). Dermatology points out the higher risk of skin cancer for people with "light eyes, red hair, fair skin, abundance of freckles," (Ward-Peterson et al., 2016, p. 1). While the biology of the skin is discussed as important, so are lifestyle and prevention, since early detection assures a higher survival rate (RKI, 2021). Medicine and healthcare relate lifestyle and prevention to different social categories. More precisely, the behavior of people and the resulting probabilities of getting skin cancer are discussed based on the categories of gender, race, class, and sexual orientation (Jiang et al., 2015; Marks et al., 2020; RKI, 2021; Stallings, 2020; Ward-Peterson et al., 2016).

Some of these categories are also discussed with regard to their biological differences, e.g. publications provide details about which group of people will develop skin cancer on which body part: men often on the trunk, women often on the extremities, ethnic minority population often on palms and soles (Cormier et al., 2006; RKI, 2021). This is seen as relevant since some parts of the body can be easier checked by oneself as others and carcinoma on some parts are better known to dermatologists than others, e.g. palms and soles constitute "unusual anatomic sites" for skin cancer by the field and therefore, are checked less (Cormier et al., 2006, p. 1907). Regarding biological aspects, skin constitution and hormone levels are discussed as having an influence and genders are appointed to these differences: the skin type of men is discussed as more easily hurt by UV light, while higher estrogen levels in women are considered to positively influence the outcome of skin cancer therapy (Stallings, 2020). While people with fairer skin seem to be at higher risk, high melanin production might be a protecting factors for people with darker skin tones (Ward-Peterson et al., 2016).

However, with regard to the survival rate of skin cancer, the field finds itself caught in a paradox: Even if white people are much more likely to get skin cancer[37], African Americans have "a 1.48-fold higher rate of risk-adjusted, stage-specific mortality

[37] The average annual age-adjusted melanoma incidence of 18.4/100,000 people for white people in the US compared to 0.8/100,000 in African American people (Cormier et al., 2006).

compared with whites" (Cormier et al., 2006, p. 1907). This paradox leads to ample discussions not only about skin-type, but also about ethnicity and race as it relates to behavioral aspects.

In general, social aspects of categories are discussed mostly based on behaviors of laypeople and of experts: e.g. POC populations are more likely to not participate in regular check-ups. Also, people who are more likely to be exposed to the sun, e.g. people from lower socioeconomic backgrounds who work outside or people from high socioeconomic backgrounds who spend their time in the sun for leisure (Jiang et al., 2015). Additionally, (mis)conceptions of healthcare personnel potentially resulting in suboptimal treatment of POC are also discussed (a contested topic in the field as shown by Cormier et al., 2006). One hypothesis is that unusual places of skin cancer on the body lead to late detection, which could be a main cause for the higher risk of African Americans in the US to succumb to the disease (Cormier et al., 2006).

4.4 AI-based skin cancer detection tools

Skin cancer detection tools are used within this social context, and are developed to assist experts and laypeople in the identification of skin cancer based on pictures of moles (del Rosario et al., 2018; Esteva et al., 2017). Soon after publications in high-ranking journals discussed the superiority of newer deep-learning based tools compared to human experts, an outcry regarding racial bias followed (Esteva et al., 2017; Noor, 2020). Critics pointed out that these tools would not work on people with darker skin reproducing already existing racial health disparities within dermatology as African Americans already have higher mortality risks through melanoma (Cormier et al., 2006; Noor, 2020). Following these critical publications, data scientists analyzed the case and it became clear that images of moles on dark skin had not been part of the training data sets used for the development of the tools, resulting in the misfunctioning of the tools on people with darker skin tones (Wen et al., 2022). The goal of my feminist science and technology analysis was to find the source of this racial bias (Baumgartner & Ernst, 2023). The following section will reanalyze the case using sociology of categorization and Epstein's (2007b, 2007d) theorizations on race in medicine as framework, then offer my second case study HIV TOS as a contrast case, and end with a comparison of the usage of categories in both AI-based tools.

4.4.1 Analysis of “race” and racism in skin cancer

People organize their world in categories to make sense of them (Brubaker, 2009; Heintz, 2021b; Zerubavel, 1996), usually those already known to them, such as race or gender. Historically, the category of race has been established during colonialization and is being reproduced within systemic racism (Brubaker, 2004; Müller, 2014). People learn as part of “optical communities” to “visually perceive and thus socially construct race” (Brekhus et al., 2010, p. 7). Constructions of similarities and difference are then related to sub-categories of the actualized category (Hirschauer, 2014; Lamont & Molnár, 2002; Wimmer, 2008; Zerubavel, 1996). In the end, people are often treated differently depending on their race (Fenstermaker & West, 2001). One example of this would be that white people are treated differently by their dermatologist than African American people (Cormier et al., 2006).

Because different subtypes of people are seen as distinct, it is only one short step away from the assumption that there are also biomedical differences which can be searched for. Critical scholars like Epstein (2007b) (analyzing how race is made relevant in the inclusion-and-difference paradigm) discuss if those differences are significant and meaningful enough to be taken into account. From the analysis of scientific and popular scientific publications on skin cancer, we see that certain categories such as gender, race, sexual orientation, and socioeconomic background are given relevance in the field (Baumgartner & Ernst, 2023). This can be theorized as doing difference, since at the same time, other categories, e.g., religion or disability are neglected (Hirschauer, 2014). These categorical differences are discussed using the ontological registers of biomedicine (for the biology of the skin), culture, or the social (for socially influenced lifestyle choices of taking care of the skin). The social is viewed as influencing the biological: e.g. taking care of one’s skin reduces the risk for skin cancer. Different cognitive schemata seem to inform each other. Practices and routines for diagnosis and treatment exist based on the defining norm of white people, who are more prone to skin cancer. We can see this in the naming of “unusual anatomic sites” for skin cancer for POC (the palm and sole) and its “unusual presentation” (Cormier et al., 2006, p.1907). Additionally, publications discuss scientific knowledge but also reproduce preconceptions about groups of people: e.g. women being more familiar with taking care of their skin than men. Significantly, these differences are often discussed based on social categories rather than the biological category of skin tone or constitution of the skin, meaning that there is a focus on the social as an ontological register as well (Cormier et al., 2006; Ward-Peterson et al., 2016).

The skin of POC being protected against melanoma together with the paradoxical higher risk of African Americans to die from it leads to further scientific discussion (Cormier et al., 2006; Ward-Peterson et al., 2016). Based on Epstein's (2007b) point that differentiating for the wrong reasons between different categories might also harm people, I suggest the following hypothesis: While information about differences based on common social categories is being produced, some of the information regarding differences might have made it into common knowledge while other information might not. For example, the information that black skin is protected from UV light is likely to be better known than higher skin cancer death rates among US African Americans compared to white people (Noor, 2020). This means when the category of race is made relevant in everyday life, the practices and routines of patients and doctors focus on the differences between races in an unhealthy way for some POC (focusing predominantly on the partial fact that they are protected from skin cancer). This first leads to unequal treatment of people of different races: white people knowing they are susceptible to skin cancer and being called in for check-ups, a lot of existing information about their specific types of skin cancer, versus less knowledge about skin cancer in POC and less pressure put on them to check their skin. The consequence of the knowledge produced within this inclusion-and-difference paradigm leads to specific (social) practices, which result in the group who is biologically less prone to die from skin cancer paradoxically becoming the one with the worst prognosis due to the influence of social factors.

However, this paradoxical outcome of African Americans having a worse prognosis for melanoma might also lead to putting more effort and money into the search for differences to understand and help resolve the paradox. This could result in strengthening the inclusion-and-difference paradigm (Epstein, 2007a). Since the field to research for these differences is biomedicine, it is only a short way to try to come up with biomedical differences. These could then again be linked to social categories and naturalize them. In this specific case, publications that focus on the difference between racial subtypes naturalize race as category. Contributing to the naturalization of race as a social category also means changing the way we think about race, and making the category contingency-aversive (Epstein, 2007d; Hirschauer, 2014). Epstein might suggest that in this case, it would be important to find out where the differences are and where they end, e.g. dermatologists need to know which different types of melanoma can be found on which body sites of different groups of people and laypeople need to know that all bodies can get skin cancer and that they should go to regular checkups no matter the skin-type.

The next section analyzes how the involvement of a skin cancer detection tool influences the aforenamed issues. It will also discuss which categories are seen as relevant and how they are conceptualized in skin cancer detection tools.

4.4.2 Racial bias in skin cancer detection tools

Just after the first outcry of skin cancer detection tools being called racially biased, Wen et al. (2022) analyzed publicly available skin cancer image data sets. The reported datasets published between 2013-2021 used the categories “sex”, “age” and “body site” (Wen et al., 2022, p.e64). It is striking to note which category is missing from most datasets: Only 2.1% of the data sets explicitly mentioned which skin tones their datasets covered. When the authors speak about skin tone, they mention the “Fitzpatrick skin type scale” (a scale usually used to assess how prone a skin type is to UV-light damage) or “ethnicity” seemingly using these terms as proxies for race or skin tone (Okoji et al., 2021; Ware et al., 2020b). Wen et al. (2022, p. e64) conclude to have found a “substantial under-representation of darker skin types” in the data sets and rate them as not representational enough “for clinical real-life settings”. Thus, pointing out the problem of lack of representational datasets for people with all skin tones. They also point out that 79% of the data sets were published in the global north and refer to “geographical and ethnic disparities in digital health data” (Wen et al., 2022, p. e71).[38]

Wen et al. (2022) also mention that a classification system for skin tone would have been relevant to diversify the pictures and represent a broader range of people for the digital tools. The most famous of these classification systems is the Fitzpatrick skin type scale (Ware et al., 2020b). The assumption that moles look different on different skin-types (at least for the AI-based tool) would be an argument to make skin tone a relevant category for skin cancer, to collect data accordingly and sort it based on the FST or a similar scale for skin tones (Das, 2021; Wen et al., 2022).

[38] This conceptualization of race/ethnicity and nation seems to be a support for Brubaker’s (2009) conceptualization of race as part of race-ethnicity-nationalism.

4.4.3 How inclusive AI leads to more categorization

Classification systems such as FST seem to be important for making a problem manageable for AI (Wen et al., 2022). They might be even more relevant for developing inclusive or fair and equitable AI which aims to make a tool work for as many groups of people as possible (Baumgartner et al., 2023; Baumgartner & Ernst, 2023; Noor, 2020).

The chosen training data determines how well the tool will work for its envisioned purpose and its envisioned target user. For inclusive tools, this means the data would have to be representative and of good quality for the whole target population (Baumgartner, 2021a; Hagendorff, 2019). If skin cancer looks different on different skin tones, inclusion of diverse variants of the problem are relevant to make sure the tool would be able to solve these different iterations of a problem and would in the end provide a high-quality outcome for a diverse group of people (Baumgartner, 2021a; Baumgartner & Ernst, 2023; Das, 2021; Wen et al., 2022). To achieve this, the population would need to be stratified according to categories relevant for the problem to be solved. This would again make these very categories relevant and reproduce or change them based on the usage in the tool.

However, which are the right groups to account for?[39] The case of skin cancer detection tools shows that data scientists are aware they need to check the properties and origin of their data (Das, 2021; Wen et al., 2022). However, data must first be collected and made available. Preconceptions prevalent within the society might lead to a situation that data is not collected in the first place. My hypothesis is that the current preconception that POC are not at risk for melanoma leads to a failure in collecting pictures from POC and in seeing them as a target group for these tools, leading instead to the development of the tools exclusively for white people. The publication by Das (2021) corroborates that, (wrongly) stating that it is not relevant to account for darker skin types in these tools. If we connect the risk of POC –with African Americans being at higher risk to die from melanoma –to a tool that cannot check their skin, this could lead to these very skin detection tools increasing the already existing health disparities (Noor, 2020). Generally speaking, the case study of skin cancer tools shows that in the end, the racial bias of the field can also be found within the AI-based tool developed within the field.

[39] Or put differently: When can we “assert that a body is a body is a body?” (Epstein, 2007c, p. 10)

Thus, my hypothesis is that AI forces us to consider existing categories because many algorithmic systems need known categories to account for them. This tendency is even stronger when aiming to develop inclusive AI. This can be done by including data of known groups of people during the developing process and checking whether the tools work for all of them in a similar way, once a prototype is has been created. Calls for fairness in AI would, therefore, reify existing categories and can naturalize social categories in medicine. The question of how the category skin tone should be introduced is not an easy one to solve, because it reminds us of an uncomfortable past and current racist incidents and experiences (see also Scheuerman et al., 2020). Ironically, AI may force us to reconsider racial categories to become useful for POC (see also Brubaker 2009).

While within this first example, social categories were made relevant in the field and in the tool, the second example HIV TOS is used as a contrast case to show how categories can be made relevant differently within a field and an AI-based tool, when there are attempts to keep the social out of AI-based tools.

4.5 The contrast case HIV TOS: conceptualizing HIV in a merely biological way

HIV treatment optimization tools function as a decision support system for virologists and healthcare professionals assisting in decisions regarding antiretroviral treatment for HIV positive individuals based on their HI viruses.

One goal of the empirical research was to analyze the social arena of HIV TOS development. The analysis showed that experts of the social worlds discursively constructed people with HIV who seek treatment in different ways (Baumgartner, 2023a). The used categories seemed to create a simplistic typology of patients to bring order to related phenomena (like resistances to medication) and materialities (which tests need to be done) to be able to act upon them. Some of the categorization schemes would help in gaining a shared understanding between the social worlds to act upon a shared goal within the arena. One example of such a category is the categorization of patients by their therapy status (therapy-naïve, therapy-failure, change-of-therapy), shared across all interviewed profession groups with the goal to ensure the survival of the patients. Here, the field distinguished between patients who had never had therapy (“therapy-naïve”) which meant they could still use all the possible options, and those who did not respond to therapy ("therapy-failure") because of newly acquired resistances. In the latter case, more drastic actions had to be taken to ensure survival and the social worlds had to work together in a concerted action to reach the goal (Baumgartner, 2023a).

Some of the categories were more informal than others. Bowker and Star (2000e, p. 332), drawing on an institutionalization point of view, would call these informal classifications "small-scale seminegotiated systems" (distinguishing them from "enforced universal systems such as race classification" under apartheid in South Africa). Drawing on Hirschauer (2014), we could say that categories exist in different aggregate states: Some categories might be more informal and used to guide practices and routines, other might be part of guidelines, such as the classification scheme based on compliance types mentioned by the physician (Baumgartner, 2023a; DAH, 2005). This classification system is used by healthcare experts to categorize their patients based on their perceived capability to adhere to a regular drug intake regime.

The physician also brought up known social categories or as healthcare experts would refer to as social determinants of health, such as age, migration background, and the intersection of female gender and being in the reproductive age or being pregnant. Most of these categories were also shared in the accounts of one virologist. Some of these categories were discussed from a biological point of view, e.g., certain medications could not be used during pregnancy. The categories of 'migration background' or 'people of African descent' were used to culturalize deficiencies like inadequate language skills or to point out different types of values or lifestyles, such as being prone to gossiping, or having a wish to give birth naturally or to breastfeed (Baumgartner, 2023a). This is relevant for the field, since breastfeeding and vaginal birth poses a risk for the baby to get infected during birth (DAIG, 2020). Language problems are a challenge for the doctor-patient relationship, since they make communication more difficult and uncertain and might increase the risk of ineffective treatment. Also, a fear of gossip means that people with HIV will not share possibly stigmatizing information within their communities and therefore lack access to already available information. The discursive constructions of the patients show how different social worlds or professional groups think about them and which information they see as relevant. It shows us how they make sense of their world and how they handle the shared goal of ensuring the survival of chronically ill patients (Baumgartner, 2023a).

Different ontological registers are used, biological ones (pregnancy) as well as social and cultural ones (being of African descent). Choice of therapy and the outcome of the therapy outcome are discussed to be influenced by lifestyles of the patient and how reliable their medication intake is (DAH, 2005). Different lifestyles of people are then linked to social categories (Baumgartner, 2023a).

4.6 HIV-treatment optimization tools

During the analysis, it soon became clear that the developed tools were simplistic with regard to data input and data output. The information used as training data and as input for the tools was only a very small amount of the data and information available on the patients, the virus, and the illness. This is even more true if compared to the information experts would generally consider when coming to a decision about a treatment regime. The training data for HIV TOS constitutes the information about which HI virus is susceptible to which active ingredient. To query an HIV TOS for a therapy suggestion, only information about HI viral RNA from a HIV positive person is needed as input, and active ingredient suggestions against the specific viruses are obtained as output (Baumgartner, 2023a).

This showed that the aforementioned discursive constructions were not taken up into the developed tools, but the categories in the tool were only biological and chemical information about the virus and reaction of ingredients to different viruses as numbers. Developers thought of genetic data as the main data to be considered for the tools and did not include any social data (Baumgartner, 2023a). This makes sense when thinking of virology – like medicine – as a field that privileges biological and or measurable data over other types of data to solve problems; Even more so as they describe their tools as examples of precision medicine, which is among other things interested in understanding a mechanistic logic of a medical problem and solving it based on biological data produced in labs (Baumgartner, 2021b).[40] The question remains as to whether the availability of more data, and especially more digital data, about social aspects would also increase its usage in digital tools.

The situational analysis could also show both what the tools are used for in the field and the knowledge that finds its way into the tools, because it had been regarded as relevant enough to solve a certain problem. One lesson learned from this analysis, is

[40] Epstein (2007a) would probably applaud the field using only relevant categories in the tool.

that the way experts speak about a problem can be very different from its depiction in a digital tool. HIV TOS are used for a very particular question and solve only a small part of the issue, while most of the problem is still solved by human experts. The tools are usually used by virologists to suggest active ingredients to physicians. The actual decision about the suggestion of treatment is then taken by the physician in consultation with the patient (Baumgartner, 2023a). This means the physician can then still take into consideration the remaining social factors. Apart from differences between analogous social worlds and digital tools, it is also possible to analyze different categorization systems in different social worlds. Through the analysis of the social worlds as well as of HIV TOS, we see which discursive construction is present in other social worlds as well, and which one is restricted to one social world. Similarly, we can see which logics are present in different social worlds and which ones are more confined to one social world. The spread of a particular logic could be a sign of a more powerful social world which can influence other social worlds by providing its way of sense-making. In this way, one social world might shape epistemologies and ontologies of a whole field. This could be the case once a new digital technology is introduced in the field. New stakeholder groups such as bioinformaticians, computer scientists, or data scientists could also change epistemologies and ontologies within the field (Baumgartner, 2023a). In my current example, the logic of the bioinformaticians, seeing patients as carriers of data, could not be found in other social worlds, because their influence in the whole field was seemingly small. This was probably also because HIV TOS only involves the assistance of virologists in a very particular issue, and the final decision considering all the relevant information is still taken by the physicians.

4.7 Discussion: Comparing HIV TOS to skin cancer detection tools

The two case studies have a few similarities and many differences when it comes to the categories used in the field and in the algorithmic tools. In both cases, biological and social categories are made relevant (Baumgartner, 2023a; Baumgartner & Ernst, 2023). When it comes to the HIV TOS example, the social is neatly kept out of the tool by the developers. While for HIV treatment social categories are commonly discussed for the tool the task is conceptualized as purely biological, even if the human still poses a very important factor in the problem of HIV resistance: e.g. the fact of taking one's medicine (compliance) is crucial to a successful therapy. Still, discursive constructions of classification systems used in the field have not found their way into HIV TOS. The tool only seems to be used to assist in a specific part of the problem. In the end, physicians consider additional (social) factors to make a decision on the treatment together with the patient. The reasons why social categories are left out of the tool can vary. When the tool was developed, the bioinformaticians aimed for an explainable tool which would be recognized by the field. Thus, they mimicked the problem-solving process of the virologists who just used virological data for their assessments of susceptibility to drug exposure. Since then, social data never found its way into the tools, probably because its availability was scarce and the advantage of including it in the assessment did not outweigh the effort to include it. Together with the tool being responsible only for a part of the problem (see also Baumgartner, 2023b), it was sufficient to let the AI-based tools solve the biological part of the problem while humans dealt with the social dimensions.

As for the skin cancer detection tools, pictures of moles are analyzed and developers probably wanted to construct the problem's solution as purely biological when programming image recognition of different looking spots with Deep Learning (Baumgartner & Ernst, 2023; Esteva et al., 2017; Wen et al., 2022). However, here the social seems to get in the way. For some of the images, information about gender, age and body site is available (Wen et al., 2022). It is unclear if this information is also used during algorithmic "calculation". Still, it is interesting that this data is part of the pictures in the databases. In contrast, the seemingly obvious category of skin tone was not considered during data collection even if dermatologists made the category race and ethnicity relevant for skin cancer risks. After an outcry about racial bias in the tools, experts from the field of data science, dermatology, and STS corroborate the relevance of the category of skin tone and race/ethnicity for the related tools (Baumgartner & Ernst, 2023; Noor, 2020; Wen et al., 2022). However, in many publications, biological and social categories seem to get mixed up, e.g., race or ethnicity and FST – a scale developed to determine the risk for UV damage for different skin types – is used as a proxy for skin tone and race (Okoji et al., 2021; Ware et al., 2020b). For behavior related to social identity and society, it may make sense to relate to the social categories of race and ethnicity. People experience different treatment based on the racial subtype people ascribe to them, rather than purely based on their skin tone, and people get different information and treatments based on their race (Cormier et al., 2006; Ward-

Peterson et al., 2016). Skin tone and race seem to be so strongly entangled that they can hardly be separated from each other. Thus, it must also be advocated to introduce them in AI-based tools.

Social categories get easily naturalized in medicine. Interestingly, naturalization based on biological conceptualization of differences was less present in the HIV TOS example where the social was kept out of AI-based tools. One hypothesis is that new discursive constructions which are not based on existing categories determine the way the field thinks about HIV treatment. Regarding HIV-infection, hypotheses are that AIDS activism did a good job in destigmatizing HIV/AIDS from being connected to specific groups of people, such as gay men or drug users. In contrast to the skin cancer example where differentiation seems to have an inclusive agenda, for HIV infection, it could rather be seen as stigmatizing and not as beneficial to point out specific groups of people at risk.

Naturalization of race happens in the skin cancer example when social and biological differences are constructed between racial subcategories at the same time as social and biological aspects are connected to each other or mixed up: Publications about darker skin being protected against skin cancer or moles being present on other body parts for African Americans (Cormier et al., 2006; Ward-Peterson et al., 2016). The social is more prevalent than the biological when physicians only diagnose skin cancer in very late stages in African Americans. This leads to a higher death rate among African Americans even if they would technically have a lower risk to get skin cancer (Ward-Peterson et al., 2016).

It is feared that the skin cancer detection tool is bound to reproduce this bias (Noor, 2020). Analysis of the tools reveal available data sets as being unrepresentative of the whole population (Wen et al., 2022). To account for more inclusive AI-based tools, representative images of different skin types would have to be entered (Baumgartner & Ernst, 2023; Wen et al., 2022). This necessitates a classification system that categorizes different skin types, and known classification systems (FST) are discussed as a solution. However, those were developed for other purposes and might not be suitable to solve the problem, because they are centered on white skin and generally do not work well (Okoji et al., 2021). Alternative classifications for skin tones are discussed within the field but would most probably also be related to racial subtypes. Moreover, they remind us of racial classifications, e.g. during apartheid, that we thought long overcome, and the problem how to make these things right is difficult to discuss due to the sensitivities surrounding the topic (Scheuerman et al., 2020). In the end, the analysis shows that categories will be more rather than less relevant when it comes to AI-based tools. This is even more true when striving to develop inclusive digital health technologies. Categories are then part of new digital materialities and institutionalized in their very specific ways.

5. Conclusion

Through the analysis of the two case studies, we learn how social worlds use known and new categories to think about issues they want to solve. Only a minor part of the categories is made relevant in the tools which are analyzed. It bears the question of whether this fact will again influence the field, or if tools will just be used as small add-ons and the real discussion will continue to take place outside of them. Biomedicine, at least in the discussed tools, still tries to stick neatly to biological categories, even if these get mixed up once they are closely related to social categories. Future research should investigate if more available information on social categories forcibly means this information will be used within AI-based tools.

Generally, digital tools in medicine and healthcare incentivize the reflection on similarities and differences, even more so if the goal is to develop fair and equitable tools that work for different groups of people. These discussions can be more difficult for contested categories, such as skin tone or race because of historical events (colonialism, slavery, apartheid) and persisting structural racism. The sociology of categorization will be important in analyzing these cases and in assisting in the discussion as regards which of these categories might be taken up in digital tools, and how.

6. References

acatech, Körber Stiftung, & ZIRIUS (Eds.). (2021). *Technikradar 2021 – Stakeholderperspektiven*. https://www.acatech.de/publikation/technikradar-2021-stakeholderperspektiven/

Adamson, A. S., & Smith, A. (2018). Machine Learning and Health Care Disparities. *JAMA Dermatology*, *154*(11), 1247–1248.

Alhutter, D., Cech, F., Fischer, F., Grill, G., & Mager, A. (2020). Algorithmic Profiling of Job Seekers in Austria. How Austerity Politics Are Made Effective. *Frontiers in Big Data*, *3*(1), 1–17.

Ammicht Quinn, R. (2021). Artificial intelligence and the role of ethics. *Statistical Journal of the IAOS*, *13*(13), 1–3.

Atkinson, P. (1984). Training for certainty. *Social Science & Medicine*, *19*(9), 949–956. https://doi.org/10.1016/0277-9536(84)90324-1

Au, L. (2021). Recent scientific/intellectual movements in biomedicine. *Social Science & Medicine*, *278*, 113950. https://doi.org/10.1016/j.socscimed.2021.113950

Baecker, D. (2018). *4.0, oder, Die Lücke die der Rechner lässt*. Merve Verlag.

Bailey, M. (2016). Misogynoir in Medical Media: On Caster Semenya and R. Kelly. *Catalyst: Feminism, Theory, Technoscience*, *2*(2), 1–31. https://doi.org/10.28968/cftt.v2i2.28800

Bailey, M. (2017). Towards a Black Feminist Health Science Studies. *Catalyst: Feminism, Theory, Technoscience*, *3*(2), 1–27.

Banner, O. (2019). Technopsyence and Afro-Surrealism's Cripistemologies. *Catalyst: Feminism, Theory, Technoscience*, *5*(1), 1–29.

Barad, K. (2015). TransMaterialities. Trans*/Matter/Realities and Queer Political Imaginings. *GLQ: A Journal of Lesbian and Gay Studies*, *21*(2–3), 387–422.

Barnett, I., Torous, J., Staples, P., Sandoval, L., Keshavan, M., & Onnela, J.-P. (2018). Relapse prediction in schizophrenia through digital phenotyping: A pilot study. *Neuropsychopharmacology*, *43*, 1660–1666.

Bath, C. (2014). Diffractive Design. In N. Marsden & U. Kempf (Eds.), *Gender-UseIT* (pp. 27–36). De Gruyter Oldenbourg.

Bauman, Z. (2005). *Moderne und Ambivalenz: Das Ende der Eindeutigkeit* (M. Suhr, Trans.; 5. Auflage). Hamburger Edition.

Baumgartner, R. (2021a). Künstliche Intelligenz in der Medizin: Diskriminierung oder Fairness. In G. Bauer, M. Kechaja, S. Engelmann, & L. Haug (Eds.), *Diskriminierung und Antidiskriminierung: Beiträge aus Wissenschaft und Praxis* (pp. 147–162). transcript.

Baumgartner, R. (2021b). Precision medicine and digital phenotyping: Digital medicine's way from more data to better health. *Big Data & Society*, *8*(2), 1–12.

Baumgartner, R. (2023a). *Personalized HIV treatment: Bringing marginalized patients to the forefront with situational analysis*. *24*(2), 1–27.

Baumgartner, R. (2023b). *Precision medicine for HIV therapy: A tale of successful risk and uncertainty management?*

Baumgartner, R., Arora, P., Bath, C., Burljaev, D., Ciereszko, K., Custers, B., Ding, J., Ernst, W., Fosch-Villaronga, E., Galanos, V., Gremsl, T., Hendl, T., Kropp, C., Lenk, C., Martin, P., Mbelu, S., Morais Dos Santos Bruss, S., Napiwodzka, K., Nowak, E., … Williams, R. (2023). Fair and equitable AI in biomedical research and healthcare: Social science perspectives. *Artificial Intelligence in Medicine*, *144*, 102658. https://doi.org/10.1016/j.artmed.2023.102658

Baumgartner, R., & Ernst, W. (2023). Künstliche Intelligenz in der Medizin? Intersektionale queerfeministische Kritik und Orientierung. *GENDER – Zeitschrift für Geschlecht, Kultur und Gesellschaft*, *15*(1), 11–25. https://doi.org/10.3224/gender.v15i1.02

Beerenwinkel, N., Lengauer, T., Däumer, M., Kaiser, R., Walter, H., Korn, K., Hoffmann, D., & Selbig, J. (2003). Methods for optimizing antiviral combination therapies. *Bioinformatics*, *19*(suppl_1), i16–i25. https://doi.org/10.1093/bioinformatics/btg1001

Beerenwinkel, N., Lengauer, T., Selbig, J., Schmidt, B., Walter, H., Korn, K., Kaiser, R., & Hoffmann, D. (2001). Geno2pheno: Interpreting genotypic HIV drug resistance tests. *IEEE Intelligent Systems*, *16*(6), 35–41. https://doi.org/10.1109/5254.972080

Beerenwinkel, N., Schmidt, B., Walter, H., Kaiser, R., Lengauer, T., Hoffmann, D., Korn, K., & Selbig, J. (2002a). Diversity and complexity of HIV-1 drug resistance: A bioinformatics approach to predicting phenotype from genotype. *Proceedings of the National Academy of Sciences*, *99*(12), 8271–8276. https://doi.org/10.1073/pnas.112177799

Beerenwinkel, N., Schmidt, B., Walter, H., Kaiser, R., Lengauer, T., Hoffmann, D., Korn, K., & Selbig, J. (2002b). Diversity and complexity of HIV-1 drug resistance: A bioinformatics approach to predicting phenotype from genotype. *Proceedings of the National Academy of Sciences*, *99*(12), 8271–8276.

Beißwenger, K.-D., & Höpfner, C. (Eds.). (1993). *Aspects of AIDS and AIDS-HILFE in Germany*. Deutsche AIDS-Hilfe e.V.

Benjamin, R. (2016). Informed Refusal. Toward a Justice based Bioethics. *Science, Technology, & Human Values*, *41*(6), 967–990.

Benjamin, R. (2019). *Race After Technology: Abolitionist Tools for the New Jim Code*. Polity Press.

Berger, P., & Luckmann, T. (1969). *Die gesellschaftliche Konstruktion der Wirklichkeit. Eine Theorie der Wissenssoziologie*. Fischer.

Birk, R. H., & Samuel, G. (2020). Can digital data diagnose mental health problems? A sociological exploration of 'digital phenotyping.' *Sociology of Health & Illness*, *42*(8), 1873–1887.

Blumer, H. (1969). *Symbolic interactionism: Perspective and method*. University of California Press.

BM für Gesundheit. (2016). *Krebsfrüherkennungsuntersuchung in der gesetzlichen Krankenversicherung (§ 25 Abs. 2 SGB V)/Mammographie. Informationsblatt Nr. 315-02*. https://bundesgesundheitsministerium.de/fileadmin/Dateien/3_Downloads/M/Mammografie/Infoblatt_Krebsfrueherkennung_Mammographie.pdf

Bogner, A., Littig, B., & Menz, W. (Eds.). (2009). *Interviewing Experts*. Palgrave Macmillan UK. https://doi.org/10.1057/9780230244276

Bosančić, S., & Keller, R. (2019). Diskursive Konstruktionen. Eine Einleitung. In S. Bosančić & R. Keller (Eds.), *Diskursive Konstruktionen: Kritik, Materialität und Subjektivierung in der wissenssoziologischen Diskursforschung* (pp. 1–8). Springer Fachmedien Wiesbaden.

Bowker, G. C., & Star, S. L. (Eds.). (2000a). Introduction: To Classify Is Human. In *Sorting things out: Classification and its consequences* (1. paperback ed., 8. print, pp. 1–32). MIT Press.

Bowker, G. C., & Star, S. L. (Eds.). (2000b). Of Tuberculosis and Trajectories. In *Sorting things out: Classification and its consequences* (1. paperback ed., 8. print, pp. 165–194). MIT Press.

Bowker, G. C., & Star, S. L. (Eds.). (2000c). *Sorting things out: Classification and its consequences* (1. paperback ed., 8. print). MIT Press.

Bowker, G. C., & Star, S. L. (2000d). The Case of Race Classification and Reclassification under Apartheid. In G. C. Bowker & S. L. Star (Eds.), *Sorting things out: Classification and its consequences* (1. paperback ed., 8. print, pp. 195–226). MIT Press.

Bowker, G. C., & Star, S. L. (Eds.). (2000e). Why Classifications Matter. In *Sorting things out: Classification and its consequences* (1. paperback ed., 8. print, pp. 319–326). MIT Press.

boyd, danah, & Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society*, *15*(5), 662–649.

Brekhus, W. H., Brunsma, D. L., Platts, T., & Dua, P. (2010). On the Contributions of Cognitive Sociology to the Sociological Study of Race: Cognitive Sociology of Race. *Sociology Compass*, *4*(1), 61–76. https://doi.org/10.1111/j.1751-9020.2009.00259.x

Brietzke, E., Hawken, E. R., Idzikowski, M., Pong, J., Kennedy, S. H., & Soares, C. N. (2019). Integrating digital phenotyping in clinical characterization of individuals with mood disorders. *Neuroscience and Biobehavioral Reviews*, *104*, 223–230.

Brubaker, R. (2004). *Ethnicity without groups*. Harvard University Press.

Brubaker, R. (2009). Ethnicity, Race, and Nationalism. *Annual Review of Sociology*, *35*(1), 21–42. https://doi.org/10.1146/annurev-soc-070308-115916

Burtnett, F. (2015). Four questions for NIMH Director Thomas R. Insel. *Counseling Today*. https://ct.counseling.org/2015/03/fourquestions-for-nimh-director-thomas-r-insel/

Butler, J. (2004). *Undoing Gender*. Routledge.

Cabitza, F., Rasoini, R., & Gensini, G. F. (2017). Unintended consequences of machine learning in medicine. *JAMA – Journal of the American Medical Association*, *318*(6), 517–518.

Carboni, C., Wehrens, R., Van Der Veen, R., & De Bont, A. (2022). Conceptualizing the digitalization of healthcare work: A metaphor-based Critical Interpretive Synthesis. *Social Science & Medicine*, *292*, 114572. https://doi.org/10.1016/j.socscimed.2021.114572

Cedars-Sinai Blog. (2019). *New Recommendations for BRCA Testing: Should You Be Screened?* https:// www.cedars-sinai.org/newsroom/new-federal-guidelines-forbrca-testing-should-you-be-screened/

Cirillo, D., Catuara-Solarz, S., Morey, C., Guney, E., Subirats, L., Mellino, S., Gigante, A., Valencia, A., Rementeria, M. J., Santuccione Chadha, A., & Mavridis, N. (2020). Sex and gender differences and biases in artificial intelligence for biomedicine and healthcare. *Npj Digital Medicine*, *3*(1), 1–11.

Clarke, A. E. (1998). *Disciplining reproduction: Modernity, American life sciences, and the problems of sex*. University of California Press.

Clarke, A. E., Friese, C., & Washburn, R. (2015). Introducing situational analysis. In A. E. Clarke, C. Friese, & R. Washburn (Eds.), *Situational analysis in practice. Mapping research with Grounded Theory* (pp. 11–76). Left Coast Press.

Clarke, A. E., Friese, C., & Washburn, R. S. (2018). *Situational analysis: Grounded theory After the interpretive turn*. Sage.

Clarke, A. E., & Montini, T. (1993). The Many Faces of RU486: Tales of Situated Knowledges and Technological Contestations. *Science, Technology, & Human Values*, *18*(1), 42–78.

Clarke, A. E., Shim, J. K., Mamo, L., Fosket, J. R., & Fishman, J. R. (2003). Biomedicalization: Technoscientific transformations of health, illness, and U.S. biomedicine. *American Sociological Review*, *68*(2), 161–194.

Clarke, A. E., & Star, S. L. (2008). The Social Worlds Framework: A Theory/Methods Package. In E. Hackett, O. Amsterdamska, M. Lynch, & J. Wajcman (Eds.), *The Handbook of Science and Technology Studies* (pp. 113–137). MIT Press.

Collins, P. H. (2019). *Intersectionality as Critical Social Theory*. Duke University Press.

Comfort, N. (2018). Genetic determinism redux. *Nature*, *561*, 461–463.

Cook, T., Atkin, H., & Wilcockson, J. (2018). Participatory Research into Inclusive Practice: Improving Services for People With Long Term Neurological Conditions. *Forum Qualitative Social Research*, *19*(1), Art. 4.

Cooper, R., & Paneth, N. (2020). Will Precision Medicine Lead to a Healthier Population? *Issues in Science and Technology*, *36*(2), 64–71.

Cormier, J. N., Xing, Y., Ding, M., Lee, J. E., Mansfield, P. F., Gershenwald, J. E., Ross, M. I., & Du, X. L. (2006). Ethnic Differences Among Patients With Cutaneous Melanoma. *Archives of Internal Medicine*, *166*(17), 1907–1914. https://doi.org/10.1001/archinte.166.17.1907

Cosgrove, L., Karter, J. M., McGinley, M., & Morrill, Z. (2020). Digital Phenotyping and Digital Psychotropic Drugs: Mental Health Surveillance Tools That Threaten Human Rights. *Health and Human Rights Journal*, *22*(2), 33–39.

Crawford, K. (2013). The Hidden Biases in Big Data. *Harvard Business Review*, *April 2013*. https://hbr.org/2013/04/the-hidden-biases-in-big-data.17

Crenshaw, K. W. (1989). Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics. *University of Chicago Legal Forum*, *1989*(1), 139–167.

Criado-Perez, C. (2020). *Unsichtbare Frauen: Wie eine von Daten beherrschte Welt die Hälfte der Bevölkerung ignoriert*. btb Verlag.

Dagum, P. (2019). Digital brain biomarkers of human cognition and mood. In H. Baumeister & C. Montag (Eds.), *Digital Phenotyping and Mobile Sensing: New Developments in Psychoinformatics* (pp. 93–108). Springer Nature Switzerland.

Dagum, P., & Montag, C. (2019). Ethical considerations of digital phenotyping from the perspective of a healthcare practitioner. In H. Baumeister & C. Montag (Eds.), *Digital Phenotyping and Mobile Sensing: New Developments in Psychoinformatics* (pp. 13–29). Springer Nature Switzerland.

DAH - Deutsche AIDS-Hilfe e.V. (Ed.). (2005). Compliance: Umgang mit der HIV-Therapie. *MED-INFO: Medizinische Informationen Zu HIV Und AIDS*, *53*.

DAH - Deutsche AIDS-Hilfe e.V. (2010). Resistenzen. *MED-INFO: Medizinische Informationen Zu HIV Und AIDS*, *72*.

DAIG - Deutsche AIDS-Gesellschaft e.V. (2020a). Deutsch-Österreichische Leitlinie zur HIV-Therapie in der Schwangerschaft und bei HIV-exponierten Neugeborenen. *AWMF Online*. https://www.awmf.org/uploads/tx_szleitlinien/055-002l_S2k_HIV-Therapie-Schwangerschaft-und- HIV-exponierten_Neugeborenen_2020-10_01.pdf

DAIG - Deutsche AIDS-Gesellschaft e.V. (2020b). Deutsch-Österreichische Leitlinien zur antiretroviralen Therapie der HIV-1-Infektion. *AWMF Online*. https://www.awmf.org/uploads/tx_szleitlinien/055- 001l_Antiretrovirale_Therapie_der_HIV_Infektion__2021-06.pdf

Das, S. (2021). Automated Bias Reduction in Deep Learning Based Melanoma Diagnosis using a Semi-Supervised Algorithm. *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 1719–1726. https://doi.org/10.1109/BIBM52615.2021.9669772

del Rosario, F., Farahi, J. M., Drendel, J., Buntinx-Krieg, T., Caravaglio, J., Domozych, R., Chapman, S., Braunberger, T., Dellavalle, R. P., Norris, D. A., Fathi, R., & Alkousakis, T. (2018). Performance of a computer-aided digital dermoscopic image analyzer for melanoma detection in 1,076 pigmented skin lesion biopsies. *Journal of the American Academy of Dermatology*, *78*(5), 927-934.

Diaz-Bone, R. (n.d.). *Neue Ausschließungsdynamiken durch Big Data-generierte Unsichtbarkeiten, Inkohärenzen und ungleiche Zeitlichkeiten*.

Diaz-Bone, R. (2018). Neue Ausschließungsdynamiken durch Big Data-generierte Unsichtbarkeiten, Inkohärenzen und ungleiche Zeitlichkeiten. In D. Houben & B. Prietl (Eds.), *Datengesellschaft* (pp. 207–230). transcript Verlag. https://doi.org/10.1515/9783839439579-009

DiMaggio, P. (1997). *CULTURE AND COGNITION*.

Epstein, S. (1996). *Impure Science: AIDS, Activism, and the Politics of Knowledge*. University of California Press.

Epstein, S. (Ed.). (2007a). From the Standard Human to Niche Standardization. In S. Epstein, *Inclusion: The politics of difference in medical research* (pp. 135–154). University of Chicago press.

Epstein, S. (2007b). *Inclusion: The politics of difference in medical research*. University of Chicago press.

Epstein, S. (Ed.). (2007c). Introduction: Health Research and the Remaking of Common Sense. In S. Epstein, *Inclusion: The politics of difference in medical research* (pp. 1–16). University of Chicago press.

Epstein, S. (Ed.). (2007d). To Profile or Not to Profile: What Difference Does Race Make. In S. Epstein, *Inclusion: The politics of difference in medical research* (pp. 203–232). University of Chicago press.

Erikainen, S., & Chan, S. (2019). Contested futures: Envisioning "personalized," "stratified," and "precision" medicine. *New Genetics and Society*, *38*(3), 308–330.

Ernst, W. (2017). Emancipatory interferences with machines? *International Journal of Gender, Science and Technology*, *9*(2), 178–196.

Ernst, W. (2021). Phänomene des Werdens: Intersektionalität, Queer, Postcolonial, Diversity und Disability Studies als Orientierungen für die Medienforschung. In J. Dorer, B. Geiger, B. Hipfl, & V. Ratković (Eds.), *Handbuch Medien und Geschlecht* (pp. 1–16). Springer VS.

Esposito, E. (2014). Algorithmische Kontingenz. Der Umgang mit Unsicherheit im Web. In A. Cevolini (Ed.), *Die Ordnung des Kontingenten* (pp. 233–249). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-531-19235-2_10

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, *542*, 115–118.

European Commission, Directorate-General for Communications Networks, Content and Technology. (2020). The Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self assessment. *Publications Office*. https://data.europa.eu/doi/10.2759/002360

European Medicines Agency (Ed.). (2018, 2007). *Atripla*

(Efavirenz/Emtricitabin/Tenofovirdisoproxil): Übersicht über Atripla und Begründung für die Zulassung in der EU. https://www.ema.europa.eu/en/documents/overview/atripla-epar-medicine-overview_de.pdf

Fenstermaker, S., & West, C. (2001). „Doing Difference“ revisited. Probleme, Aussichten und der Dialog in der Geschlechterforschung. In B. Heintz (Ed.), *Geschlechtersoziologie* (pp. 236–249). Westdeutscher Verlag.

Ferryman, K., & Pitcan, M. (2018). What is precision medicine? *Data & Society*, 1–27.

Figueroa, C. A., Luo, T., Aguilera, A., & Lyles, C. R. (2021). The need for feminist intersectionality in digital health. *The Lancet Digital Health*, *3*(8), 526–533.

Findlay, P., Lindsay, C., McQuarrie, J., Bennie, M., Corcoran, E. D., & Van Der Meer, R. (2017). Employer Choice and Job Quality: Workplace Innovation, Work Redesign, and Employee Perceptions of Job Quality in a Complex Health-Care Setting. *Work and Occupations*, *44*(1), 113–136. https://doi.org/10.1177/0730888416678038

Foucault, M. (1974). *Die Ordnung des Diskurses*. Hanser.

Foucault, M. (2008). *The Birth of Biopolitics. Lectures at the Collège de France, 1978–79*. Palgrave Macmillan.

Fox, R. C. (1957). Training for uncertainty. In R. K. Merton, G. G. Reader, & P. L. Kendall (Eds.), *The Student-Physician: Introductory Studies in the Sociology of Medical Education* (pp. 207–242). Harvard University Press. https://doi.org/10.4159/harvard.9780674366831.c13

Fox, R. C. (1980). The Evolution of Medical Uncertainty. *The Milbank Memorial Fund Quarterly. Health and Society*, *58*(1), 1–49. https://doi.org/10.2307/3349705

Fröhlich, H., Balling, R., Beerenwinkel, N., Kohlbacher, O., Kumar, S., Lengauer, T., Maathuis, M. H., Moreau, Y., Murphy, S. A., Przytycka, T. M., Rebhan, M., Röst, H., Schuppert, A., Schwab, M., Spang, R., Stekhoven, D., Sun, J., Weber, A., Ziemek, D., & Zupan, B. (2018). From hype to reality: Data science enabling personalized medicine. *BMC Medicine*, *16*(1), 150. https://doi.org/10.1186/s12916-018-1122-7

Fujimura, J. H. (1987). Constructing “Doable Problems” in Cancer Research: Articulating Alignment. *Social Studies of Science*, *17*(2), 257–293.

Furlow, B. (2020). Information overload and unsustainable workloads in the era of electronic health records. *The Lancet Respiratory Medicine*, *8*(3), 243–244. https://doi.org/10.1016/S2213-2600(20)30010-2

Gallo, R. C., & Montagnier, L. (2003). The Discovery of HIV as the Cause of AIDS. *New England Journal of Medicine*, *349*(24), 2283–2285.

Garcia-Ceja, E., Riegler, M., Nordgreen, T., Jakobsen, P., Oedegaard, K. J., & Tørresen, J. (2018). Mental health monitoring with multimodal sensing and machine learning: A survey. *Pervasive and Mobile Computing*, *51*, 1–26.

G-BA - Gemeinsamer Bundesausschuss (Ed.). (2004). *HIV-Resistenztestung: Zusammenfassender Bericht des Arbeitsausschusses “Ärztliche Behandlung” des*

Bundesausschusses der Ärzte und Krankenkassen über die Beratung des Jahres 2003 zur Bewertung der HIV-Resistenztestung gemäß §135 Abs. 1 SGB V. Unterausschuss „Ärztliche Behandlung“ des Gemeinsamen Bundesausschusses.

Ghosh, A. K. (2004). Understanding medical uncertainty: A primer for physicians. *The Journal of the Association of Physicians of India*, *52*, 739–742.

Gigerenzer, G., & Weiler, S. (2020). *Google AI erkennt Brustkrebs besser als die erfahrensten Radiologen*. Deutsche Statistische Gesellschaft. https://dstatg.de/unstatistik-des-monats-januar

Gillespie, T. (2014). The Relevance of Algorithms. In T. Gillespie, P. J. Boczkowski, & K. A. Foot (Eds.), *Media Technologies* (pp. 167–194). The MIT Press. https://doi.org/10.7551/mitpress/9780262525374.003.0009

Grommé, F. (2016). Data mining 'problem youth': Looking closer But Not seeing better. In I. van der Ploeg & J. Pridmore (Eds.), *Digitizing Identities* (pp. 163–183). Routledge, Taylor & Francis.

Gutiérrez Rodríguez, E. (2011). Intersektionalität oder: Wie nicht über Rassismus sprechen? In S. Hess, N. Langreiter, & E. Timm (Eds.), *Intersektionalität Revisited. Empirische, theoretische und methodische Erkundungen* (pp. 77–100). transcript.

Hacking, I. (2005). Why race still matters. *Daedalus*, *134*(1), 102–116. https://doi.org/10.1162/0011526053124460

Hagendorff, T. (2019). Maschinelles Lernen und Diskriminierung: Probleme und Lösungsansätze. *Österreichische Zeitschrift für Soziologie*, *44*(S1), 53–66. https://doi.org/10.1007/s11614-019-00347-2

Halford, S., Obstfelder, A., & Lotherington, A.-T. (2010). Changing the record: The interprofessional, subjective and embodied effects of electronic patient records: Changing the record. *New Technology, Work and Employment*, *25*(3), 210–222. https://doi.org/10.1111/j.1468-005X.2010.00249.x

Heimstädt, M., Egbert, S., & Esposito, E. (2021). A Pandemic of Prediction: On the Circulation of Contagion Models between Public Health and Public Safety. *Sociologica*, 1-24 Pages. https://doi.org/10.6092/ISSN.1971-8853/11470

Heintz, B. (2010). Numerische Differenz. Überlegungen zu einer Soziologie des (quantitativen) Vergleichs / Numerical Difference. Toward a Sociology of (Quantitative) Comparisons. *Zeitschrift für Soziologie*, *39*(3), 162–181. https://doi.org/10.1515/zfsoz-2010-0301

Heintz, B. (2017). Kategoriale Ungleichheit und die Anerkennung von Differenz. In S. Hirschauer (Ed.), *Un/doing Differences* (pp. 79–116). Velbrück Wissenschaft. https://doi.org/10.5771/9783845292540-80

Heintz, B. (2021a). Big Observation – Ein Vergleich moderner Beobachtungsformate am Beispiel von amtlicher Statistik und Recommendersystemen. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, *73*(S1), 137–167. https://doi.org/10.1007/s11577-021-00744-0

Heintz, B. (2021b). Kategorisieren, Vergleichen, Bewerten und Quantifizieren im Spiegel sozialer Beobachtungsformate. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, *73*(S1), 5–47. https://doi.org/10.1007/s11577-021-00741-3

Helsana. (2017). *Brustkrebs-Früherkennung durch Mammographie-Screening*. https://helsana.ch/de/blog/koerper/koerperwissen/brustkrebs-frueherkennung.html?articleSource=gesundheitskompetenz

Hengstenberger, M. (2005). *Gynäkologie von Frau zu Frau*. Springer.

Hirschauer, S. (2014). Un/doing Differences. Die Kontingenz sozialer Zugehçrigkeiten. *Zeitschrift für Soziologie*, 22.

Hirschauer, S. (2021). Menschen unterscheiden. Grundlinien einer Theorie der Humandifferenzierung. *Zeitschrift für Soziologie*, *50*(3–4), 155–174. https://doi.org/10.1515/zfsoz-2021-0012

Hofmann, B., & Svenaeus, F. (2018). How medical technologies shape the experience of illness. *Life Sciences, Society and Policy*, *14*(1), 3. https://doi.org/10.1186/s40504-018-0069-y

Hombach, S. (2018). *Medizingeschichte: Vom Aderlass bis zur brusterhaltenden Therapie*. https://amoena.com/be-de/amoena-life/medizingeschichte-vom-aderlass-bis-zur-brusterhaltenden-therapie/

Hood, C., & Rothstein, H. (2001). Risk Regulation Under Pressure. *Administration & Society*, *33*(1), 21–53.

Hood, L., Balling, R., & Auffray, C. (2012). Revolutionizing medicine in the 21st century through systems approaches. *Biotechnology Journal*, *7*(8), 992–1001.

Hopman, R. (2020). Opening up forensic DNA phenotyping: The logics of accuracy, commonality and valuing. *New Genetics and Society*, *39*(4), 424–440.

Horn, W. (2001). AI in medicine on its way from knowledge-intensive to data-intensive systems. *Artificial Intelligence in Medicine*, *23*(1), 5–12. https://doi.org/10.1016/S0933-3657(01)00072-0

Horsley, N. (2021). Intellectual autonomy after artificial intelligence: The future of memory institutions and historical research. In N. Lushetich (Ed.), *Big data: A new medium?* (pp. 1–15). Routledge.

Houben, D., Prietl, B., & Deutsche Gesellschaft für Soziologie (Eds.). (2018). *Datengesellschaft: Einsichten in die Datafizierung des Sozialen*. Tagung "DatenGesellschaft!?“, Bielefeld. Transcript.

Huckvale, K., Venkatesh, S., & Christensen, H. (2019). Toward clinical digital phenotyping: A timely opportunity to consider purpose, quality, and safety. *Npj Digital Medicine*, *2*(1), 1–11.

Jacob, E. K. (2004). Classification and Categorization: A Difference that Makes a Difference. *Library Trends*, *52*(3), 515–540.

Jacobs, M., He, J., Pradier, M. F., Lam, B., Ahn, A. C., McCoy, T. H., Perlis, R. H., Doshi-Velez, F., & Gajos, K. Z. (2021). Designing AI for trust and collaboration in time-constrained medical decisions: A sociotechnical lens. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–14. https://doi.org/10.1145/3411764.3445385

Jain, S. H., Powers, B. W., Hawkins, J. B., & Brownstein, J. S. (2015). The digital phenotype. *Nature Biotechnology*, *33*, 462–463.

Jenkins, R. (1997). *Rethinking ethnicity: Arguments and explorations*. Sage.

Jermutus, E., Kneale, D., Thomas, J., & Michie, S. (2022). Influences on User Trust in Healthcare Artificial Intelligence: A Systematic Review. *Wellcome Open Research*, *7*, 65. https://doi.org/10.12688/wellcomeopenres.17550.1

Jiang, A. J., Rambhatla, P. V., & Eide, M. J. (2015). Socioeconomic and lifestyle factors and melanoma: A systematic review. *British Journal of Dermatology*, *172*(4), 885–915. https://doi.org/10.1111/bjd.13500

Keller, E. F. (2002). *The Century of the Gene*. Harvard University Press.

Keller, R., Hirseland, A., Schneider, W., & Viehöfer, W. (Eds.). (2005). *Die diskursive Konstruktion von Wirklichkeit*. UVK.

Kim, K., Heinze, K., Xu, J., Kurtz, M., Park, H., Foradori, M., & Nolan, M. T. (2018). Theories of health care decision making at the end of life: A meta-ethnography. *Western Journal of Nursing Research*, *40*(12), 1861–1884.

Kitchin, R. (2014a). Big Data, new epistemologies and paradigm shifts. *Big Data & Society*, *1*(1), 205395171452848. https://doi.org/10.1177/2053951714528481

Kitchin, R. (2014b). Big data, new epistemologies and paradigm shifts. *Big Data and Society*, *1*(1), 1–12.

Kitchin, R. (2017). Thinking critically about and researching algorithms. *Information, Communication & Society*, *20*(1), 14–29. https://doi.org/10.1080/1369118X.2016.1154087

Klein, H. K., & Kleinman, D. L. (2002). The Social Construction of Technology: Structural Considerations. *Science, Technology, & Human Values*, *27*(1), 28–52. https://doi.org/10.1177/016224390202700102

Klinger, C. (2013). Überkreuzende Identitäten – Ineinandergreifende Strukturen. Plädoyer für einen Kurswechsel in der Intersektionalitätsdebatte. In C. Klinger & G.-A. Knapp (Eds.), *ÜberKreuzungen. Fremdheit, Ungleichheit, Differenz* (pp. 38–67). Westfälisches Dampfboot.

Klinger, C. (2022). Life Care/Lebenssorge and the Fourth Industrial Revolution. In H. Nagl-Docekal & W. Zacharasiewicz (Eds.), *Artificial Intelligence and Human Enhancement: Affirmative and Critical Approaches in the Humanities* (pp. 101–124). De Gruyter.

Klöckner, J., & Rybicki, B. (2020, October 21). Datenspende für bessere Früherkennung: KI wird im Kampf gegen Brustkrebs immer wichtiger. *Handelsblatt*. https://handelsblatt.com/technik/digitale-revolution/digitale-revolution-daten-spende-fuer-bessere-frueherkennung-ki-wird-im-kampf-gegen-brustkrebs-immer-wichtiger/26288516.html?ticket=ST-3286334-MtdH32bJcbUdIQcVA2it-ap2

Kumari, S., Chouhan, U., & Suryawanshi, S. K. (2017). Machine learning approaches to study HIV / AIDS. *Bioscience Biotechnology Research Communications*, *10*(1), 34–43.

Kuria, E. N. (2014). Theorizing Race(ism) While NeuroGendering. In S. Schmitz & G. Höppner (Eds.), *Gendered Neurocultures: Feminist and Queer Perspectives on Current Brain Discourses* (pp. 109–123). Zaglossus.

Lamont, M., & Molnár, V. (2002). The Study of Boundaries in the Social Sciences. *Annual Review of Sociology*, *28*(1), 167–195. https://doi.org/10.1146/annurev.soc.28.110601.141107

Lee, S. M. (1993). Racial classifications in the US census: 1890–1990. *Ethnic and Racial Studies*, *16*(1), 75–94. https://doi.org/10.1080/01419870.1993.9993773

Lenz, S. (2021). "More like a support tool": Ambivalences around digital health from medical developers' perspective. *Big Data & Society*, *8*(1), 205395172199673. https://doi.org/10.1177/2053951721996733

Liang, Y., Zheng, X., & Zeng, D. D. (2019). A survey on big data-driven digital phenotyping of mental health. *Information Fusion*, *52*, 290–307.

Light, D. Jr. (1979). Uncertainty and control in professional training. *Journal of Health and Social Behavior*, *20*(4), 310–322. https://doi.org/10.2307/2955407

Liu, T. F., & Shafer, R. W. (2006). Web resources for HIV type 1 genotypic-resistance test interpretation. *Clinical Infectious Diseases*, *42*(11), 1608–1618. https://doi.org/10.1086/503914

Loi, M. (2019). The digital phenotype: A philosophical and ethical exploration. *Philosophy & Technology*, *32*, 155–171.

Lorde, A. (2000). *Auf Leben und Tod. Krebstagebuch*. Fischer Taschenbuch.

Low, S. K., Zembutsu, H., & Nakamura, Y. (2018). Breast cancer: The translation of big genomic data to cancer precision medicine. *Cancer Science*, *109*(3), 497–506.

Lücking, P. (2020). Automatisierte Ungleichheit. Wie algorithmische Entscheidungssysteme gesellschaftliche Machtverhältnisse (re)produzieren. In netzforma* e.V. (Ed.), *Wenn KI, Dann Feministisch. Impulse auf Wissenschaft und Aktivismus* (pp. 65–76).

Lupton, D. (2015). *Digital Sociology*. Routledge.

Lupton, D. (2016). *The Quantified Self*. Polity Press.

Lupton, D. (2018a). *Digital Health – Critical and Cross-Disciplinary Perspectives*. Routledge.

Lupton, D. (2018b). How do data come to matter? Living and becoming with personal data. *Big Data & Society*, *5*(2), 1–11.

Lupton, D., & Maslen, S. (2017). Telemedicine and the senses: A review. *Sociology of Health & Illness*, *39*(8), 1557–1571. https://doi.org/10.1111/1467-9566.12617

Lux, M. P., Emons, J., Bani, M. R., Wunderle, M., Sell, C., Preuss, C., Rauh, C., Jud, S. M., Heindl, F., Langemann, H., Geyer, T., Brandl, A.-L., Hack, C. C., Adler, W., Schulz-Wendtland, R., Beckmann, M. W., Fasching, P. A., & Gass, P. (2019). Diagnostic Accuracy of Breast Medical Tactile Examiners (MTEs): A Prospective Pilot Study. *Breast Care – Multidisciplinary Journal for Research, Diagnosis and Therapy*, *14*, 41–47.

Lyon, D. (2003). *Surveillance as Social Sorting*. Routledge.

Mager, A., & Mayer, K. (2019). Body data—Data body: Tracing ambiguous trajectories of data bodies between empowerment and social control in the context of health. *Momentum Quarterly – Zeitschrift Für Sozialen Fortschritt*, *8*(2), 95–109.

Maiers, C. (2017). Analytics in action: Users and predictive data in the neonatal intensive care unit. *Information, Communication & Society*, *20*(6), 915–929. https://doi.org/10.1080/1369118X.2017.1291701

Marks, D. H., Arron, S. T., & Mansh, M. (2020). Skin Cancer and Skin Cancer Risk Factors in Sexual and Gender Minorities. *Dermatologic Clinics*, *38*(2), 209–218.

Mau, S. (2017). *Das Metrische Wir. Über Die Quantifizierung Des Sozialen*. Suhrkamp.

McCoy, L. (2020). What makes medical knowledge: On the epistemological tensions between "evidence-based medicine" and "personalized medicine." *University of Toronto Medical Journal*, *97*(1), 27–31.

McKinney, S. M., Sieniek, M., Godbole, V., Godwin, J., Antropova, N., Ashrafian, H., Back, T., Chesus, M., Corrado, G. S., Darzi, A., Etemadi, M., Garcia-Vicente, F., Gilbert, F. J., Halling-Brown, M., Hassabis, D., Jansen, S., Karthikesalingam, A., Kelly, C. J., King, D., … Shetty, S. (2020). International evaluation of an AI system for breast cancer screening. *Nature*, *577*, 89–94.

McLean, S., Ressler, K., Chase Koenen, K., & et al. (2020). The AURORA study: A longitudinal, multimodal library of brain biology and function after traumatic stress exposure. *Molecular Psychiatry*, *25*, 283–296.

Melcher, J., Hays, R., & Torous, J. (2020). Digital phenotyping for mental health of college students: A clinical review. *Evidence-Based Mental Health*, *23*(4), 161–166.

Mennicken, A., & Espeland, W. N. (2019). What's new with numbers? Sociological approaches to the study of quantification. *Annual Review of Sociology*, *45*, 223–245.

Metaxiotis, K. S., & Samouilidis, J. E. (2000). Expert systems in medicine: Academic illusion or real power? *Information Management and Computer Security*, *8*(2–3), 75–79.

Metz, R. (2018). *The smartphone app that can tell you're depressed before you know it yourself*. https://www.technologyreview.com/s/612266/the-smartphone-app-thatcan-tell-youre-depressed-before-you-know-it-yourself/

Meuser, M., & Nagel, U. (2010). ExpertInneninterview. Zur Rekonstruktion spezialisierten

Sonderwissens. In R. Becker & B. Kortendiek (Eds.), *Handbuch für Frauen- und Geschlechterforschung* (pp. 376–379). VS Verlag für Sozialwissenschaften.

Milne, R. (2020). The rare and the common: Scale and the genetic imaginary in Alzheimer's disease drug development. *New Genetics and Society*, *39*(1), 101–126.

Mindstrong Health. (2021). *Using science to help us seek the truth*. https://mindstrong.com/science/

Mol, A. (2008). *The logic of care*. Routledge.

Müller, M. (2014). „The evils of racism and the wealth of diversity" – Zum Bedeutungswandel der Rassenkategorie bei den UN-Weltkonferenzen gegen Rassismus: "The evils of racism and the wealth of diversity": Semantic Changes in the Category of Race in UN World Conferences against Racism. *Zeitschrift für Soziologie*, *43*(6), 402–420. https://doi.org/10.1515/zfsoz-2014-0602

Nardini, C., Annoni, M., & Schiavone, G. (2012). Mechanistic understanding in clinical practice: Complementing evidence-based medicine with personalized medicine. *Journal of Evaluation in Clinical Practice*, *18*(5), 1000–1005. https://doi.org/10.1111/j.1365-2753.2012.01907.x

Nassehi, A. (2019). *Muster: Theorie der digitalen Gesellschaft*. C.H. Beck.

National Human Genome Research Institute. (n.d.). *The CLIA Framework*. https://www.genome.gov/Pages/PolicyEthics/GeneticTesting/The_CLIA_Framework.pdf

Nielsen, E. (2019). *Disrupting Breast Cancer Narratives: Stories of Rage and Repair*. University of Toronto Press.

Noble, S. U. (2018). *Algorithms of oppression. How search engines reinforce racism*. New York University Press.

Noor, P. (2020). Can we trust AI not to further embed racial bias and prejudice? *BMJ*, *368*, 10–12.

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, *366*(6464), 447–453.

Offenberger, U. (2019). Anselm Strauss, Adele Clarke und die feministische Gretchenfrage. Zum Verhältnis von Grounded-Theory-Methodologie und Situationsanalyse. *Forum Qualitative Social Research*, *20*(2), Art. 6.

Okoji, U. K., Taylor, S. C., & Lipoff, J. B. (2021). Equity in skin typing: Why it is time to replace the Fitzpatrick scale. *British Journal of Dermatology*, *185*(1), 198–199. https://doi.org/10.1111/bjd.19932

O'Neil, C. (2017). *Angriff der Algorithmen. Wie sie Wahlen manipulieren, Berufschancen zerstören und unsere Gesundheit gefährden*. Carl Hanser Verlag.

Oudshoorn, N. (2008). Diagnosis at a distance: The invisible work of patients and healthcare professionals in cardiac telemonitoring technology. *Sociology of Health & Illness*, *30*(2), 272–288. https://doi.org/10.1111/j.1467-9566.2007.01032.x

Ozaydin, B., Hardin, J. M., & Chhieng, D. C. (2016). Data Mining and Clinical Decision Support Systems. In E. S. Berner (Ed.), *Clinical Decision Support Systems* (pp. 45–68). Springer International Publishing. https://doi.org/10.1007/978-3-319-31913-1_3

Panch, T., Mattie, H., & Celi, L. A. (2019). The “inconvenient truth” about AI in healthcare. *Npj Digital Medicine*, *2*(1), 77. https://doi.org/10.1038/s41746-019-0155-4

Parsons, T. (2005). *The social system*. Taylor & Francis.

Pasquale, F. (2016). *The black box society: The secret algorithms that control money and information* (First Harvard University Press paperback edition). Harvard University Press.

Peters, T. (2012). *Playing God? Genetic Determinsm and Human Freedom*. Routledge.

Poferl, A. (2004). *Die Kosmopolitik des Alltags*. edition sigma.

Polanyi, M. (2009). *The tacit dimension* (22nd ed.). University of Chicago Press.

Pols, J. (2010). The Heart of the Matter. About Good Nursing and Telecare. *Health Care Analysis*, *18*(4), 374–388. https://doi.org/10.1007/s10728-009-0140-1

Porter, T. M. (2020). *The rise of statistical thinking, 1820-1900* (New edition). Princeton University Press.

Prainsack, B. (2017a). *Personalized medicine: Empowered patients in the 21st century*. New York University Press.

Prainsack, B. (2017b). *Personalized medicine. Empowered patients in the 21st century?* New York University Press.

Prietl, B. (2019). Die Versprechen von Big Data im Spiegel feministischer Rationalitätskritik. *GENDER*, *2019*(3), 11–25.

Queertactics. (2021). *The Cancer Journals revisited*. https://queer-tactics.at/the-cancer-journals-revisited-lentos/

Rammert, W. (Ed.). (1995). *Soziologie und künstliche Intelligenz: Produkte und Probleme einer Hochtechnologie*. Campus Verlag.

Rammert, W. (Ed.). (1998). *Wissensmaschinen: Soziale Konstruktion eines technischen Mediums ; das Beispiel Expertensysteme*. Campus-Verl.

Rammert, W. (2000). *Nicht-explizites Wissen in Soziologie und Sozionik: Ein kursorischer Überblick* (pp. 147–166). Technische Universität Berlin. https://doi.org/10.1007/978-3-531-90485-6_9

Ranjan, Y., Rashid, Z., Stewart, C., Conde, P., Begale, M., Verbeeck, D., Boettcher, S., Dobson, R., & Folarin, A. (2019). RADAR-base: Open source mobile health platform for collecting, monitoring, and analyzing data using sensors, wearables, and mobile devices. *JMIR mHealth and UHealth*, *7*(8), 1–12.

Reed, K., Kochetkova, I., & Whitby, E. (2016). Visualising uncertainty: Examining women’s views on the role of Magnetic Resonance Imaging (MRI) in late pregnancy. *Social Science & Medicine*, *164*, 19–26. https://doi.org/10.1016/j.socscimed.2016.07.012

Reiss, J., & Sprenger, J. (2020). Scientific objectivity. *Stanford Encyclopedia of Philosophy*, *Winter 2020*. https://plato.stanford.edu/archives/win2020/entries/scientific-objectivity/

Rhee, S.-Y. (2003). Human immunodeficiency virus reverse transcriptase and protease sequence database. *Nucleic Acids Research*, *31*(1), 298–303. https://doi.org/10.1093/nar/gkg100

RKI - Robert Koch Institut. (2021). HIV in Deutschland 2020 | Erfassung der SARS-CoV-2-PCR-Testzahlen. *Epidemiologisches Bulletin*, *47*. https://www.rki.de/DE/Content/Infekt/EpidBull/Archiv/2021/Ausgaben/47_21.pdf?__blob=publicationFile

RKI - Robert Koch Institut, & Zentrum für Krebsregisterdaten. (2021). Malignes Melanom der Haut. In RKI - Robert Koch Institut & Zentrum für Krebsregisterdaten (Eds.), *Krebs in Deutschland* (pp. 72–75). Robert Koch Institut. https://krebsdaten.de/Krebs/DE/Content/Krebsarten/ Melanom/melanom_node.html

Roberts, D. E., & Rollins, O. (2020). Why Sociology Matters to Race and Biosocial Science. *Annual Review of Sociology*, *46*(1), 195–214. https://doi.org/10.1146/annurev-soc-121919-054903

Ruckenstein, M., & Schüll, N. D. (2017). The datafication of health. *Annual Review of Anthropology*, *46*, 261–278.

Rüppel, J. (2019). "Now is a time for optimism": The politics of personalized medicine in mental health research. *Science Technology and Human Values*, *44*(4), 581–611.

Saeb, S., Lattie, E. G., Schueller, S. M., Kording, K. P., & Mohr, D. C. (2016). The relationship between mobile location sensor data and depression symptom severity. *PeerJ*, *4*, e2537.

Saukko, P. (2018). Digital health – a new medical cosmology? The case of 23andMe online genetic testing platform. *Sociology of Health and Illness*, *40*(8), 1312–1326.

Scamell, M., & Stewart, M. (2014). Time, risk and midwife practice: The vaginal examination. *Health, Risk & Society*, *16*(1), 84–100. https://doi.org/10.1080/13698575.2013.874549

Scheuerman, M. K., Wade, K., Lustig, C., & Brubaker, J. R. (2020). How We've Taught Algorithms to See Identity: Constructing Race and Gender in Image Databases for Facial Analysis. *Proceedings of the ACM on Human-Computer Interaction*, *4*(CSCW1), 1–35. https://doi.org/10.1145/3392866

Schmid, S. (2018). The Discovery of HIV-1. *Nature Milestones*, *4*. https://www.nature.com/articles/d42859-018-00003-x

Schneider, I. (2021). Diskriminierungsgefahren und Regulationsansätze bei der medizinischen Nutzung von Künstlicher Intelligenz. *Zeitschrift Für Medizinische Ethik*, *67*(3), 327–350.

Schubert, C. (2007). Risk and safety in the operating theater: An ethnographic study of sociotechnical practices. In R. V. Burri & J. Dumit (Eds.), *Biomedicine as culture: Instrumental practices, technoscientific knowledge, and new modes of life* (pp. 123–138). Routledge.

Schubert, C. (2015). Situating technological and societal futures. Pragmatist engagements with computer simulations and social dynamics. *Technology in Society*, *40*, 4–13. https://doi.org/10.1016/j.techsoc.2014.07.003

Schubert, C. (2022). Die langen Schatten der Digitalisierung in der Medizin. Zur Entwicklung und Nutzung digitaler Technologien aus sozialwissenschaftlicher Sicht. In A. Manzei-Gorsky, C. Schubert, & J. Von Hayek (Eds.), *Digitalisierung und Gesundheit* (pp. 67–94). Nomos Verlagsgesellschaft mbH & Co. KG. https://doi.org/10.5771/9783748922933-67

Schulz, M., & Zinn, J. O. (2023). Rationales of risk and uncertainty and their epistemological foundation by new phenomenology. *Journal of Risk Research*, *26*(3), 219–232. https://doi.org/10.1080/13669877.2022.2162105

Schütz, A., & Luckmann, T. (1994). *Strukturen der Lebenswelt* (5. Aufl). Suhrkamp.

Schwennesen, N. (2019). Algorithmic assemblages of care: Imaginaries, epistemologies and repair work. *Sociology of Health & Illness*, *41*(S1), 176–192. https://doi.org/10.1111/1467-9566.12900

Seely, A. J. E. (2013). Embracing the Certainty of Uncertainty: Implications for Health Care and Research. *Perspectives in Biology and Medicine*, *56*(1), 65–77. https://doi.org/10.1353/pbm.2013.0009

Seyfert, R. (Ed.). (2017). *Algorithmuskulturen: Über die rechnerische Konstruktion der Wirklichkeit*. Algorithmic Cultures (Conference), Bielefeld. Transcript.

Shafer, R. W. (2006). Rationale and Uses of a Public HIV Drug-Resistance Database. *The Journal of Infectious Diseases*, *194*(s1), S51–S58. https://doi.org/10.1086/505356

Shortliffe, E. H. (2019). Artificial Intelligence in Medicine: Weighing the Accomplishments, Hype, and Promise. *Yearbook of Medical Informatics*, *28*(01), 257–262. https://doi.org/10.1055/s-0039-1677891

Sjoding, M. W., Dickson, R. P., Iwashyna, T. J., Gay, S. E., & Valley, T. S. (2020). Racial Bias in Pulse Oximetry Measurement. *New England Journal of Medicine*, *383*(25), 2477–2478.

Spinnewijn, L., Aarts, J., Verschuur, S., Braat, D., Gerrits, T., & Scheele, F. (2020). Knowing what the patient wants: A hospital ethnography studying physician culture in shared decision making in the Netherlands. *BMJ Open*, *10*, e032921.

Stallings, A. F. (2020). *Why more men than women get skin cancer*. https://westchesterhealth.com/blog/why-more-men-than-women-get-skin-cancer/

Star, S. L. (1985). Scientific work and uncertainty. *Social Studies of Science*, *15*(3), 391–427.

Stockinger, G. (2002, April 7). Katastrophe für die Frauen. *Der Spiegel*.

https://spiegel.de/wissenschaft/katastrophe-fuer-die-frauen-a-387efb77-0002-0001-0000-000022019408

Strauss, A., & Corbin, J. (1994). Grounded theory methodology: An overview. In *Handbook of qualitative research.* (pp. 273–285). Sage Publications, Inc.

Strauss, A. L. (1978). A Social World Perspective. *Studies in Symbolic Interaction*, *1*, 119–128.

Strauss, A. L. (1984). Social Worlds and Their Segmentation Processes. *Studies in Symbolic Interaction*, *5*, 123–139.

Strauss, A. L. (1993). *Continual permutations of action*. Aldine de Gruyter.

Strauss, A. L., & Corbin, J. (1994). Grounded theory methodology: An overview. In N. K. Denzin (Ed.), *Handbook of qualitative research* (pp. 273–285). Sage.

Suchman, L. (2019). Feministische Science & Technology Studies (STS) und die Wissenschaften vom Künstlichen. *GENDER*, *2019*(3), 56–83.

Tajfel, H. (1978). *Differentiation between social groups: Studies in the social psychology of intergroup relations.* (pp. xv, 474). Academic Press.

Terhart, H. (2014). *Körper und Migration*. transcript.

Timmermans, S., & Angell, A. (2001). Evidence-based medicine, clinical uncertainty, and learning to doctor. *Journal of Health and Social Behavior*, *24*(4), 342–359.

Timmermans, S., & Berg, M. (2003). The practice of medical technology. *Sociology of Health & Illness*, *25*(3), 97–114. https://doi.org/10.1111/1467-9566.00342

Topol, E. J. (2019a). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, *25*(1), 44–56. https://doi.org/10.1038/s41591-018-0300-7

Topol, E. J. (2019b). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, *25*, 44–56.

Torous, J., Kiang, M. V., Lorme, J., & Onnela, J.-P. (2016). New tools for new research in psychiatry: A scalable and customizable platform to empower data driven smartphone research. *JMIR Mental Health*, *3*(2), e16.

U.S. Government. (2021). *Electronic Code of Federal Regulations: Part 493 – Laboratory Requirements*. https://www.ecfr.gov/current/title-42/chapter-IV/subchapter-G/part-493

van Dijk, J. (2014). Datafication, dataism and dataveillance: Big data between scientific paradigm and ideology. *Surveillance and Society*, *12*(2), 197–208.

Vienna online. (2022). *„SkinScreener": Früherkennung von Hautkrebs mittels App*. https://vienna.at/skinscreener-frueherkennung-von-hautkrebs-mittels-app/7359523

Wachter, S., Mittelstadt, B., & Russel, C. (2021). Why Fairness Cannot Be Automated:

Bridging the Gap Between EU Non-Discrimination Law and AI. *Computer Law & Security Review*, *41*, 105567.

Ward-Peterson, M., Acuña, J. M., Alkhalifah, M. K., Nasiri, A. M., Al-Akeel, E. S., Alkhaldi, T. M., Dawari, S. A., & Aldaham, S. A. (2016). Association Between Race/Ethnicity and Survival of Melanoma Patients in the United States Over 3 Decades: A Secondary Analysis of SEER Data. *Medicine*, *95*(17), 1–6.

Ware, O. R., Dawson, J. E., Shinohara, M. M., & Taylor, S. C. (2020a). Racial Limitations of Fitzpatrick Skin Type. *Cutis*, *105*(2), 77–80.

Ware, O. R., Dawson, J. E., Shinohara, M. M., & Taylor, S. C. (2020b). Racial Limitations of Fitzpatrick Skin Type. *SKIN OF COLOR*, *105*(2), 77–80.

Weiner, K., Martin, P., Richards, M., & Tutton, R. (2017). Have we seen the geneticisation of society? Expectations and evidence. *Sociology of Health and Illness*, *39*(7), 989–1004.

Weiss, K. M. (2017). Is precision medicine possible? *Issues in Science and Technology*, *34*(1).

Wen, D., Khan, S. M., Ji Xu, A., Ibrahim, H., Smith, L., Caballero, J., Zepeda, L., De Blas Perez, C., Denniston, A. K., Liu, X., & Matin, R. N. (2022). Characteristics of publicly available skin cancer image datasets: A systematic review. *The Lancet Digital Health*, *4*(1), e64–e74.

Wiggert, K. (2021). *The role of scenarios in scripting (the use of) medical technology: The case of data-driven clinical decision support systems*. Technische Universität Berlin. https://doi.org/10.14279/DEPOSITONCE-11441

Wimmer, A. (2008). The Making and Unmaking of Ethnic Boundaries: A Multilevel Process Theory. *American Journal of Sociology*, *113*(4), 970–1022. https://doi.org/10.1086/522803

Winter, P., & Carusi, A. (2022). 'If you're going to trust the machine, then that trust has got to be based on something': Validation and the co-constitution of trust in developing artificial intelligence (AI) for the early diagnosis of pulmonary hypertension (PH). *Science & Technology Studies*, 58–77. https://doi.org/10.23987/sts.102198

Yang, X. J., Wickens, C. D., & Hölttä-Otto, K. (2016). How users adjust trust in automation: Contrast effect and hindsight bias. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *60*(1), 196–200. https://doi.org/10.1177/1541931213601044

Zazzi, M., Cozzi-Lepri, A., & Prosperi, M. C. F. (2016). Computer-aided optimization of combined anti-retroviral therapy for HIV: New drugs, new drug targets and drug resistance. *Current HIV Research*, *14*(2), 101–109. https://doi.org/10.2174/1570162X13666151029102254

Zerubavel, E. (1996). Lumping and splitting: Notes on social classification. *Sociological Forum*, *11*(3), 421–433. https://doi.org/10.1007/BF02408386

Zinn, J. O. (Ed.). (2008). *Social theories of risk and uncertainty: An introduction*. Blackwell Pub.

Zinn, J. O. (2016). 'In-between' and other reasonable ways to deal with risk and uncertainty: A review article. *Health, Risk & Society*, *18*(7–8), 348–366. https://doi.org/10.1080/13698575.2016.1269879

Zinn, J. O., & Olofsson, A. (2019). Introduction. In A. Olofsson & J. O. Zinn (Eds.), *Researching Risk and Uncertainty: Methodologies, Methods and Research Strategies* (1st ed. 2019, pp. 1–28). Springer International Publishing: Imprint: Palgrave Macmillan. https://doi.org/10.1007/978-3-319-95852-1