AI & SOCIETY
https://doi.org/10.1007/s00146-023-01825-8

OPEN FORUM

Perspectives of patients and clinicians on big data and AI in health: a comparative empirical investigation

Patrik Hummel[1] · Matthias Braun[2] · Serena Bischoff[3] · David Samhammer[3] · Katharina Seitz[4] · Peter A. Fasching[4] · Peter Dabrock[3]

Received: 22 December 2022 / Accepted: 16 November 2023

Abstract

Background Big data and AI applications now play a major role in many health contexts. Much research has already been conducted on ethical and social challenges associated with these technologies. Likewise, there are already some studies that investigate empirically which values and attitudes play a role in connection with their design and implementation. What is still in its infancy, however, is the comparative investigation of the perspectives of different stakeholders.

Methods To explore this issue in a multi-faceted manner, we conducted semi-structured interviews as well as focus group discussions with patients and clinicians. These empirical methods were used to gather interviewee's views on the opportunities and challenges of medical AI and other data-intensive applications.

Results Different clinician and patient groups are exposed to medical AI to differing degrees. Interviewees expect and demand that the purposes of data processing accord with patient preferences, and that data are put to effective use to generate social value. One central result is the shared tendency of clinicians and patients to maintain individualistic ascriptions of responsibility for clinical outcomes.

Conclusions Medical AI and the proliferation of data with import for health-related inferences shape and partially reconfigure stakeholder expectations of how these technologies relate to the decision-making of human agents. Intuitions about individual responsibility for clinical outcomes could eventually be disrupted by the increasing sophistication of data-intensive and AI-driven clinical tools. Besides individual responsibility, systemic governance will be key to promote alignment with stakeholder expectations in AI-driven and data-intensive health settings.

Keywords Big data · Health · Stakeholder · Interviews · Ethics · Privacy · Self-determination

Abbreviations

CTG Cardiotocogram
EHR Electronic Health Record(s)

1 Introduction

Intensified and accelerated datafication and automation are receiving increased attention in health contexts. The health sector is being transformed by "the processing of large quantities of data, with the aim of discerning patterns and thus gaining novel insights" (Council et al. 2017) as well as expansion and acceleration along various dimensions such as the value, volume, velocity, variety, veracity, and variability of data (Andreu-Perez et al. 2015). The resulting complexity is navigated by applications to process big data, such as machine learning or artificial intelligence (AI) more generally (Yu et al. 2018), often as part of a promise to drive the personalization of health services (Suwinski et al. 2019). In the presence of these powerful tools, commentators begin to wonder what exactly is and will be the role of the human clinician vis-à-vis the machine (Coiera 2018; Braun et al. 2020), e.g., whether the former eventually be replaced at

✉ Patrik Hummel
p.a.hummel@tue.nl

[1] Philosophy and Ethics Group, Department of Industrial Engineering and Innovation Sciences, TU Eindhoven, De Rondom 70, 5612 AP Eindhoven, The Netherlands

[2] Institute for Social Ethics, Faculty of Protestant Theology, University Bonn, Am Hof 1, 53113 Bonn, Germany

[3] Institute for Systematic Theology, Chair of Systematic Theology II (Ethics), Friedrich-Alexander-Universtität Erlangen-Nürnberg, Kochstraße 6, 91054 Erlangen, Germany

[4] Universitätsklinikum Erlangen, Frauenklinik Universitätsstraße 21/23, 91054 Erlangen, Germany

Published online: 07 January 2024

least partially (Darcy et al. 2016; Goldhahn et al. 2018) by the latter, or whether a confluence of both (Topol 2019) is the most fruitful way of thinking about their relation in the long-term.

At the same time, an increasing number of more cautionary takes highlight gaps between the enthusiasm and attention levels around these technologies and the evidence about their efficacy. Despite the hype (Chen and Asch 2017; Maughan 2017; Emanuel and Wachter 2019), reality checks of ambitious claims (Chin-Yee and Upshur 2019; Wilkinson et al. 2020) and more concrete, successful use cases are needed to demonstrate how beyond bold visions and buzz words, data-intensive and automated tools advance biomedical research and clinical care effectively. A systematic review published in March 2020 found only 10 records for deep learning randomized clinical trials, only two of which have been published. Moreover, "[o]f 81 non-randomised clinical trials identified, only nine were prospective and just six were tested in a real world clinical setting" (Nagendran et al. 2020). Trial and reporting guidelines (Liu et al. 2020; Rivera et al. 2020) for assessing AI-driven medicine are just beginning to emerge. Moreover, data-driven and automated tools are typically devised to carry out tasks that are highly specific, addressing particular snapshots of clinical practice. In line with this, the testing and refinement of such tools typically takes place in relatively confined and almost artificial scenarios in which particular conceptions of system performance, e.g., in terms of sensitivity and specificity of diagnostic classification, are compared with the performance of human clinicians (e.g., Gulshan et al. 2016). Less attention is devoted to the actual *implementation* challenges (Shaw et al. 2019) of AI and big-data-based tools more generally in real-world clinical workflows and the broader downstream effects of deploying these systems, once deemed sufficiently effective in testing scenarios. There is a clear need for cross-disciplinary research to examine the full range of effects such introduction has on clinical practice, health outcomes, and the self-conception and various sets of stakeholder expectations.

Besides the gap between hype and reality, a further issue of concern in connection with health-related big data is the extensive need for patient data for the development and refinement of tools ideally advancing research, care, and the common good on the one hand, and possible privacy issues on the other. Health data are intimate and sensitive data, and threats to privacy can be aggravated by big data and machine learning (Price and Cohen 2019). Additional points of contention emerge from the fact that ever expanding amounts of health-related data are processed by large private-sector organizations, raising shifts and tensions in the purposing of such data "[w]hen digital health meets digital capitalism" (Sharon 2018). Different routes and emphases are defended to harness medical big data and without compromising the rights and interests of data subjects. Some caution against focusing solely on informed consent of data subjects and argue that issues of non-exploitation and fair risk-benefits distributions should be the focus of attention (McCoy et al. 2020). Potentially compatible with this approach, others demand the alignment of technological infrastructures and regulatory environments for enabling patients and researchers to retain control over their data and maintain data sovereignty (Hummel et al. 2018, 2021).

To make progress on these issues, the perspectives of patients and medical experts using and deploying big data applications are a distinctive source of knowledge. They can contribute first-hand judgements of potentials and challenges in harnessing health-related big data. These stakeholders operate and interact in the context of clinical workflows that are increasingly shaped and transformed by these new technologies. At the same time, neither patients nor clinicians are homogenous user groups, but vary in backgrounds, expectations, expertise, responsibilities, and embeddings into clinical workflows and organizational structures of one or more healthcare institutions. Each of these aspects raises distinctive aspects for further consideration throughout the process from conceiving of new data-intensive tools to implementing and continuously reviewing them in clinical routine. Different disciplines have thus begun to engage in systematic empirical examinations of the perspectives of researchers and clinical practitioners on digital health (Fernau et al. 2018; Gabriels and Moerenhout 2018; Fiske et al. 2020; Martani et al. 2021; Martinho et al. 2021; Jongsma et al. 2021; Sand et al. 2022). The number of studies investigating the perspectives of patients on this topic is more limited (Richter et al. 2019; Richardson et al. 2021; Köngeter et al. 2022).

The aim of the present interview study is to tap into this source of evidence and to shed light on the perspectives of patients and clinicians on big data applications in the clinic. The distinctive contribution of our study is that it gathers and compares attitudes and perceptions articulated by patients and clinicians, respectively. It does so on the basis of querying patients and clinicians of one healthcare institution, the Women's Hospital Erlangen. Their perspectives are sought to address the following research questions: how do patients and clinicians perceive potentials and challenges posed by big data in the clinic? Which forms of big-data-driven health-related inferences are deemed desirable or problematic, and on what grounds? And which shifts in clinical practice and physician–patient interaction do they anticipate? These questions were investigated by means of semi-structured interviews and group discussions, which took as their starting points the concrete stakeholder experiences with health-related big data.

2 Methods

Our study consists of two interview series: one to explore patient perspectives on clinical big data methods in which 40 current, recent, or potential patients were interviewed. In the second interview series, 15 clinicians were interviewed. Both interview series share as a common starting point that they were driven by prompts and stimuli relating to participants' personal experiences with health-related big data: a set of cases presented to the patients, and questions concerning the relevance of big data methods in the professional environment of clinicians.

To explore patient perspectives, 24 patients of the Women's Hospital were interviewed. The sample was supplemented with 16 volunteers who are not currently seeking care. For the selection of interviewees, the goal was to sample a variety of patients from different age groups and disease backgrounds (see Table 1):

- Group P-I: currently or previously under cancer treatment;
- Group P-II: not currently or previously under cancer treatment.

Participants were assigned to focus group discussions, grouped by age (< 40 years, 40 < years), whether or not they have or had cancer, and educational background as far as this was logistically feasible. With the onset of the COVID-19 pandemic, we switched to one-on-one telephone interviews. Focus group discussions lasted between 60 and 90 min, telephone interviews around 30 min.

The focus group discussions followed a semi-structured interview guide, and a condensed version of the same guide was used for the telephone interviews. Semi-structured interviews and group discussions were chosen as a method to investigate the research questions since they unite a consistent frame across interviews to allow comparability on the one hand with the flexibility to attend to specific points that are raised in a particular group discussion or interview on the other. One methodological difficulty we anticipated was that participants would have varying levels of background knowledge and familiarity with the technologies on which our study was to explore their perspectives. To this end, we chose to include three brief case studies in the interview guide. Case studies are a research method for capturing and analyzing data pertaining research questions by reference to specific instances of a subject matter, especially in connection with research that investigates 'how'- or 'why'-questions (Yin 2018). Mindful of the fact that neither case studies (Baškarada 2014) nor their application in qualitative empirical research (Diefenbach 2008) are without methodological challenges, we used cases with the rather minimal purpose of providing input and a common reference point for discussion based on which interviewees were invited to lay out their thoughts. Presentation of these cases was limited to descriptions at a relatively granular level of detail, outlining only the main ideas behind the design of the respective system or technology, and purposefully eluded any normative implications as far as possible.

Table 1 Patient interviewees

	≤ 40 years	> 40 years
Group P-I	5	15
Group P-II	12	8

- Case 1: Google Flu Trends as a method to trace and to predict flu infections (Ginsberg et al. 2009).
- Case 2: Machine learning programs as a method to diagnose mental health issues on the basis of a user's Instagram profile (Reece and Danforth 2017).
- Case 3: Big data methods for personalized cancer treatment (van 't Veer and Bernards 2008; Thangue and Kerr 2011).

After a brief input (a PowerPoint slide each on the first two cases and a brief news video on the third), participants were asked about their views on these technologies. Towards the end of the interview, a set of more general questions on the processing of broadly health-related data were raised.

To explore clinician perspectives, 15 one-on-one interviews were conducted in the Women's Hospital Erlangen. In line with methodological literature on expert interviews (Meuser et al. 2009), we designed an interview guide with open-ended questions, starting with cursory explorations of their professional self-conception, e.g., whether the focus is on clinical care, medical research, or both, and perceptions of their own role in the Women's Hospital. The guide then turned to big-data-related items, the extent to which big data shapes or could/will affect clinicians' work environment, and the roles they take on in data-intensive research and care. Formulation of the interview questions were informed by ongoing debates in the literature and exploratory meetings with leadership at the study site. For the selection of interviewees, the goal was to sample individuals with different levels of experience and positions across the organizational structure of one and the same institution. The roster of clinicians was divided into three career stages:

- Assistant clinicians in years 1–3 of their training;
- Assistant clinicians in years 4–5 of their training, medical specialists (Facharzt/Fachärztin), and deputy heads of department (stellvertretende(r) Oberarzt/Oberärztin);

- Heads of department, leadership positions.

Individuals from each group were drawn randomly and asked about their interest and willingness to participate in the study. Five individuals per group were identified for inclusion.

We were also interested in whether clinicians' perspectives differ depending on familiarity with and routine deployment of big data methods. Thus, after each interview, the interviewer placed the interviewee in one of the following three ex post categories to reflect the individuals' proximity or distance to clinical big data methods in their clinical workflows:

- Group C-1: no or very little interaction with big data methods;
- Group C-2: occasional interaction with big data methods, familiarity with potential use cases;
- Group C-3: daily or regular interaction with big data methods and automated systems.

The resulting distribution of these two groupings is shown in Table 2.

Both the patient and clinician interviews were recorded, transcribed, and analyzed using MAXQDA and Atlas.ti. Any information that could relate to specific individuals was removed before the analysis. In the analysis, we followed techniques from qualitative content analysis (Mayring 2014), specifically from Grounded Theory (Glaser and Strauss 1967) and applications of these approaches to the qualitative analysis of interviews (Kruse 2015). In planning and carrying out the present study as well as preparing the manuscript, we followed the methodological guidance of Berthelsen et al. (Berthelsen et al. 2018) who draw together and systematize various streams in Grounded Theory. We created a basic coding scheme that drew from the questions and themes of the interview guides and was refined iteratively out of the material itself to arrive at a systematic account of the qualitative content. The coding was divided into several phases and the scheme was differentiated and adapted in constant discussion between three coders. With this inductive coding, we searched for connections between certain kinds of statement and speakers' roles, socio-demographic backgrounds, and other features (Mayring 2019) and compared different codes in the qualitative material (Böhm 2019) to guide the analysis.

One key component of both interview sets is the term 'big data'. While in the patient interviews we sketched illustrative cases, we did not pre-specify a definition of 'big data' in the clinician interviews and remained largely agnostic on specific understandings. Whenever interviewees were queried about the meaning of this term by the patients or clinicians, they suggested the working definitions mentioned in *Introduction*, but generally, the goal was to learn as much as possible on participants' understandings of the term without priming them in favor of particular definitions.

3 Results

We present the results for the two interview series clustered by the groupings just described, i.e., disease status for the patient group, and proximity to clinical big data methods for the clinicians. From respective prompts and stimuli just outlined, we characterized the following four recurring domains inductively (Thomas 2006; Chandra and Shang 2019):

– Concrete experiences with clinical big data,
– Perceived benefits and potentials of clinical big data that were mentioned,
– Perceived risks and challenges,
– Strategies that could help to navigate the field.

In square brackets, we provide references to the *Supplementary File* which lists relevant direct quotations from the German interview transcripts. The English direct quotations in the present text are our translations.

3.1 Patients

We first sketch observations that were distinctively raised by Groups P-I and P-II, respectively, before turning to recurring themes in both Groups.

3.1.1 Group P-I: current or previous cancer treatment

Individuals from Group P-I drew connections to their own case histories to illustrate potentials of Case 3, i.e.,

Table 2 Clinician interviewees

	Total no. of individuals (female, male)	Assist. clinicians (yrs. 1–3)	Assist. clinicians (yrs. 4–5), specialists, deputy heads	Heads of department, leadership
Group C-1	4 (3, 1)	1 (1, 0)	4 (4, 0)	0
Group C-2	8 (3, 5)	2 (0, 2)	1 (0, 1)	4 (1, 3)
Group C-3	3 (1, 2)	2 (1, 1)	0	1 (0, 1)

data-intensive cancer research and care to better understand and address causes of cancer [P1], risk management, treatment choice, and patient experience [P2]. The hope was that collecting and especially connecting as well as making accessible larger amounts of clinical data could help to mitigate burdens for patients [P3]. For example, individuals in Group P-I reported about repeated and, in their view, redundant testing and measuring of health parameters across different healthcare sites and even different units of one and the same hospital, both throughout treatment regimens and within single hospital visits. At the level of research, big data methods were expected to eventually extend and advance the refinement of treatments that some of the interviewees described as having saved their lives [P4].

Challenges were anticipated when the reliability of big-data-based predictions is not obvious. For example, it was anticipated that whenever a patient is faced with a prediction about an increased risk for cancer or recurrence [P5-6], this could cause anxiety, doubt, and actually elevate rather than mitigate subjective uncertainty. Another class of concerns was that when processing personal genomic data for cancer research, pharmaceutical industry might benefit disproportionately compared to patients [P7]. Finally, perceived risks loom if intimate genomic data is considered outside the clinic, and as a result constrains the freedom and well-being of the patient, e.g., if it were to complicate changing one's health insurance [P8] or cause stigma or exploitation [P9]. Interviewees expressed "fear" of such scenarios but noted that open and transparent consultations with their clinician helped to contextualize risks and appreciate the benefits of data-intensive genomic research [P12].

Indeed, while appreciating benefits, participants also pointed to limitations of purely data-based methods, highlighting the significance of the role of their clinician. One particular example were gene expression tests (Wallden et al. 2015) for which some patients perceived an ambiguity [P10] in how these tests shaped the work of their treating clinicians: on the one hand, these tests are seen as operating somewhat autonomously in the sense that the prognostic information is wholly machine-driven, arising from a fully automated analysis of the tumor; "this is a machine […], the clinician did nothing" [P10]. On the other hand, these prognostic outputs are just "bare figures" [P10] that are dependent on interpretation by a human clinician to be meaningful.

3.1.2 Group P-II: no current or previous cancer treatment

While individuals from Group P-II spoke as observers and in more distanced and hypothetical terms about Case 3, they gave similar reasons as Group P-I for the perceived utility of harnessing large sets of genomic data, i.e., to enhance treatment effectiveness through individualization [P13-14]. Intensified datafication in health infrastructures could further facilitate the pooling of knowledge and competence between and across different research institutions and clinical sites. Ideally, each cancer patient is then diverted to places that are able to treat their particular disease profile in accordance with the most recent and highest quality evidence [P17]. Interviewees anticipated that accelerations along these lines will lend "hope" to current and future patients [P15]. Amongst the concerns was that attempts to generate large amounts of data for research will lead to the testing of at least some interventions that are futile, unnecessary, or more invasive than necessary [P16]. Another reservation was about the possibility that even methods that consider very large amounts of clinical data might still elide or incompletely consider "so many aspects that the computer does not know" and which yet contribute to disease or are relevant to ideal treatment choice [P18].

3.1.3 Recurring themes across both Groups

Across both Groups, there was a clear tendency of participants in their views of the three Cases: participants were generally amenable to and appreciated the potential benefits of personalized cancer treatments. However, they were more skeptical of allowing and facilitating health-related inferences on the basis of social media data.

A variety of reasons for such skepticism were provided. First, participants were less convinced that the tools from cases 1 and 2 work reliably and actually generate benefits. For example, they suspected Google search queries, and even the symptoms being googled, are in fact non-specific to the inferred disease [P19]. A system that tries to draw inferences about flu waves (hypothetically, as Google Flu Trends was discontinued in 2011) will thus track irrelevant features or patterns. This is a stark contrast to the case of personalized cancer care that has more concrete tangible benefits and is directly relevant to the experiences and expectations of participants, particularly Group P-I.

Second, a perceived asymmetry amongst the Cases is the question of who benefits from data processing and in what way. All Cases concern data processing that does not only affect one's own self-interest, but also shapes the interests and care of others. The methods or algorithms at issue have the potential to continuously learn on the basis of user or patient data. In this way, they become more powerful and informative about individuals not initially involved in the processing. Such dynamics were deemed in principle acceptable and commendable provided that certain conditions are in place. One condition relates to the purpose of processing. For example, "as long as cancer research and the curing of cancer is the goal, this is not a problem" [P26]. In contrast, Cases 1 and 2 involve private-sector organizations as data processors, and a recurring theme was that commercialization negatively affects desirability: "As soon as this [Google

Flu Trends] goes in the direction of moneymaking, I do not approve of it" [P21]. Overall, specification of the goals for which processing is carried out, and of the beneficiaries of such processing, are highly relevant to acceptability.

Third, the Cases differ in terms of their potential harms. The reservations about Case 3 (in addition to those just mentioned) largely concern the personal nature of the genomic data and related potentials for misuse and data security breaches. Due to the personal nature of genomic data, these are certainly non-negligible, but also continuous with concerns about just any kind of sensitive data. Cases 1 and 2 induce further kinds of concerns: besides the mentioned risk of poor forecasting and diagnosis, there was a clear perception that Cases 1 and 2 operate behind users' backs by drawing health-related inferences without users' awareness and explicit consent. The primary purpose for users to make their data available is usage of the platform, not facilitating health-related inferences, and they are typically not aware of the secondary analyses. Indeed, none of the interviewees indicated familiarity with the possibility of the kinds of methods from Cases 1 and 2. While Case 1 was deemed primarily surprising and as raising general questions about the range of health-related inferences a service provider can draw about users, Case 2 elevated this concern even further by specifically inferring mental health issues for which, regrettably, many perceive or anticipate stigma.

Fourth, the foregoing observation that there are acceptable versus more problematic forms of health-related inferences points towards the issue of *controlling* when, how, by whom, and for what purposes such inferences are drawn. On this issue, perspectives diverged in their details, but had in common that Cases 1 and 2 undercut such control in some way. According to one stream of argument, if data are processed for the benefit of others, consent remains important and data subjects should retain control over data access while remaining informed about the purposes of their processing. In line with this, the vast majority of interviewees expressed that they consider their patient records 'their' data, i.e., these data belong to and are in some sense owned by them, which is why the respective individual is entitled to control access.

In addition, some interviewees maintained, mostly in connection with Case 3, that there are certain conditions under which it is conceivable to dispense with consent requirements for processing an individual's data: if the processing results in social value and benefits patients rather than industry and/or fully anonymity is guaranteed in the sense that no connections can possibly be drawn between the data subject and their data, then it can be acceptable to proceed without the individual's consent. As one extension of this view about the potential social value to be generated from health-related data through biomedical research, some interviewees maintained that in Case 3, genomic data obtained by analyzing their tumor belong not the individual, but to the research community. While this kind of argument has the appearance of justifying the softening of control and consent, it must be highlighted that such softening is deemed acceptable against the backdrop of conditions related to purpose and anonymity. In other words, less control is deemed acceptable only because and to the extent that the processing actually complies with the individual's preferences.

Other interviewees unpacked a claim that is primarily descriptive rather than evaluative, but relevant to the status of the issue of control: Cases 1 and 2 de facto rest on modes and practices of consenting that are quite different from Case 3. Participants explained that mere usage of platforms, social media, and other services in which they engage in their everyday lives already facilitate such inferences. Such usage entails a more or less conscious decision of the user to make data available to the processor, including for secondary processing with health-related import. By virtue of posting on social media or using a search engine, one has already agreed to such processing; "the moment I post, I have consented" [P20]. The alternative is refraining from usage of such services altogether. This implicit kind of consent is decidedly different from clinical research, e.g., in Case 3, where explicit consent mechanisms and purpose limitation in accordance with the consent given are practiced and expected by data subjects [P22].

Again, this point has the appearance of justifying the kind of processing in Cases 1 and 2 as instances of implicit consent. However, the picture becomes less clear once an additional perspective is considered. To begin with, as just described (under 'Third'), Cases 1 and 2 indicate that it can become somewhat opaque to the user what happens with their data and which health-related inferences are drawn or feasible, rendering it a challenge to generate *informed* consent. Moreover, especially in connection with Google Flu Trends, many responded with what in the analysis process we termed 'fatalism', i.e., the attitude that such process are unavoidable and inescapable: "But don't they have the data anyways? " [P11]. Moreover, in describing skepticism and reservations about such putatively unavoidable data processing and health-related inferences further, some participants invoked worries about the "surveillance state" [P23, P24], insurance companies, and specifically expressed concerns about data processors located in China. The unifying feature of these concerns seems to be that there are grave power and knowledge asymmetries between data subjects and processors. These give rise to hope, but also skepticism and diffuse unease, each of them entangled with affective and rational components and more or less well-specified conditions to retain control over how data affects their lives.

3.2 Clinicians

We present our results from the clinician interviews grouped by proximity to big data applications (see *Methods*). Interviews began with exploring interviewees' notions and perceptions of big data in the clinic. Some interviews quickly turned to definitional issues of what counts as a big data application, and all of them discussed potentials and challenges.

3.2.1 Group C-1: little to no interaction with big data methods

The availability and consideration of large amounts of data will advance research that benefits from increasingly large cohorts, e.g., in registry trials, large number of data points per participant, and the construction of databases to facilitate research and data sharing. One clinician involved in research specifically pointed to the potential of large clinical and biomaterial data bases that could facilitate targeted enquiries to generate and to test hypotheses. Clinicians focusing on clinical care were hopeful that big data could broaden the evidence base, provide a better basis for treatment decisions, and ultimately put clinicians in a position to choose the best therapy with greatest certainty [C1].

As another potential benefit, interviewees saw a clear potential to use big data and automated methods as a safety mechanism to track and to evaluate proposed treatment decisions. Clinicians can make mistakes, and a system monitoring proposed or planned interventions could help to flag errors [C2].

As for the challenges mentioned, big data leads to an abundance of information that needs to be filtered and contextualized to guide practice. On the one hand, such abundance can in principle facilitate patient involvement and empowerment, e.g., when medical information is accessible to them online. On the other hand, separating sound from misleading evidence can be a serious challenge for patients, clinicians, and their interactions in the course of shared decision-making [C3].

Somewhat in contrast to the abovementioned hype around clinical big data, clinicians were particularly enthusiastic about a more foundational transformation that would go hand in hand with using big data methods: the introduction and roll-out of electronic health records (EHR). While EHR are already the norm in many other countries, paper-based patient records remain common in German hospitals, although there are ongoing initiatives to systematically roll out EHR (Stegemann and Gersch 2021). Clinicians were hopeful that EHR could allow the clinician to access the patient case history, past treatments, and diagnostic tests more easily, and thereby facilitate more efficient decision-making processes. One pragmatic challenge was the transitioning from at least partially paper-based records towards new forms of datafication as this process will result in additional workloads when converting existing records into digital formats [C5]. These points exemplify a tendency already apparent in Group C-1 and emphasized further by the other Groups: while advantages and potentials of big data are recognized and appear promising, these potentials have not yet been realized in the *status quo* in clinical care. Interviewees phrased the generation, accessibility, and processing of biomedical big data as a counterfactual. EHR, let alone advanced forms of automation that raise the hype and enthusiasm (see *Introduction*) are portrayed as being largely absent from day-to-day care and service provision [C4].

3.2.2 Group C-2: occasional interaction with big data methods

Clinicians whom we classified as interacting with big data occasionally also highlighted intra- and inter-institutional information flows as a major potential: effectively exchanging information with other hospitals and units, e.g., to consider general practitioner and cardiological data alongside oncological data to weigh treatment options with as much information at hand as possible, and minimizing redundant diagnostic interventions. Likewise for recruitment into research projects, it is helpful to know the patient's case history in detail, e.g., previous treatments and details on tumor biology.

Participants from Group C-2 provided a set of observations about how harnessing larger amounts of clinical data could facilitate the automation of certain parts of clinical decision-making processes. One line of statements in Group C-2 coheres with the foregoing observation to some extent: when speaking about the impact of big data on their ongoing work, interviewees echoed an asymmetry between research and care, with big data already being relevant, useful, and transformative in research but not yet shaping clinical care. It was pointed out that big data could facilitate useful forms of automation and even self-learning, AI-driven decision support tools, but this was phrased as a potential development, not an actuality.

Another somewhat contrasting line of statements suggested that there are already forms of clinical, data-driven automation that shape service provision. Participants from Group C-2 raised two concrete examples. First, one respondent reported on a big data tool integrated into a molecular tumor board for breast cancer. Based on genetic test results of a given patient, the tool points clinicians' attention towards potential treatment regimens. Second, another interviewee, after first remarking that big data does not yet affect routine workflows in clinical care, indicated that setting aside machine learning or other cutting-edge information technologies, upon a closer look there are in fact already

plenty of data-intensive automated tools in the clinic. The interviewee highlighted the application Oxford CTG as an example: it processes intrapartum fetal heartrate and detects non-obvious but clinically relevant short-term variations (Dawes et al. 1996; Pardey et al. 2002). To identify deviations, it uses a reference database containing over 100.000 nonstress test readings [C9]. These examples appeared to undercut the assertion that methods relying on high-volume data processing are still largely removed from everyday clinical care.

Whether already in place or not, interviewees distinguished a range of potential or actual added value of applications that combine large amounts of data with automated decision support. First, as shown in the tumor board example, such tools could be used to identify and suggest potential avenues for treatment, thereby introducing new or reassuring input for consideration into decision-making processes. Ideally, machine-driven outputs are presented for consideration to the human clinician; "it is good if it runs automatically, and then goes to a clinician who checks and says 'yes, this works' and with this, we save a lot of time" [C8]. If so, automatization could improve workflows while the clinician ensures safety and human control. This was portrayed as being compatible with remaining committed to fundamental differences between these tools and human experts: "when it comes to decisions on therapies, these are actually taken in tumor board conferences where all specialists are present" [C27], and "structures are not at this point yet where we can replace the expert" [C27] with a neural network. Second, interviewees argued that such tools could be seen as analogues to medical guidelines, unifying and making accessible the state of the art that is ordinarily reflected in guidelines. Third, such tools could be seen as analogues to experienced experts, reflecting and mimicking their wealth of knowledge, thereby harnessing evidence bases that no or very few human experts can leverage. Fourth, automated tools were once again appreciated as potential safety mechanisms that could intervene in situations where human decision-makers can make mistakes [C6].

Limitations were highlighted as well. Like the patient groups, clinicians acknowledged that data-intensive tools not only mitigate but sometimes also elevate uncertainty. First, within the abundance of available information, it becomes more difficult to separate sound hypotheses from unsupported claims. Second, if technologies are opaque and obscure which inputs matter in which ways, feelings of uncertainty and diminished confidence result for both clinicians and patients [C7]. Given that clinicians have a better understanding of the technology under consideration, this introduces additional informational asymmetries and potential skepticism and ambivalence into the patient-clinician interaction. Third, apparently in line with points raised in the patient interviews as well, clinicians too identified certain roles and skills that big data methods are unable to take over as a matter of principle. For example, speaking with regard to neonatal heartbeat as assessed by the Oxford CTG, one interviewee distinguished the question of how a measurement relates to evidence-based notions of normalcy on the one hand, and which intervention is warranted in light of this determination on the other; "there are two different levels, right? One is the calculation of the parameter itself and the other is how this parameter really is associated with an outcome" [C10]. An automated system can assist in assessing the former, but human judgement inevitably comes in when assessing relevant implications for the latter, predicting effects, and taking decisions on this basis.

3.2.3 Group C-3: regular interaction with big data methods

Clinicians whom we classified as interacting with big data methods on a regular basis outlined several points that overlap or resonate with the other Groups. In research, big data was seen as enabling the expansion of cohorts in studies, e.g., in observational trials of increasing size. In clinical care, big data would enable increases in the number of available data points per individual, in particular longitudinal data from the first visit up until regular follow-ups post-treatment [C11-14, C30]. Moreover, when coupled with hospital IT systems, EHR could facilitate systematic quality assurance and harnessing of data to guide decision-making in the individual patient encounter as well as at the organizational level [C15-16]. Regarding possible challenges, data security and privacy breaches were highlighted [C19]. There is uncertainty about the extent to which such data, even if hidden or anonymized today, will be available or inferable to third parties in the future, and also which conclusions third parties could draw from them on particular individuals. Health and occupational disablement insurance could discriminate against individuals with high-risk profiles identified through big data analyses [C20, C25]. In some countries, accessibility of health data could also enable state agencies to discriminate against individuals with certain case histories [C26]. Strategies to handle such risks were outlined as well, ranging from technical designs to structural factors, e.g., strictly delimiting access rights to patient data to entrusted healthcare institutions [C21–22].

Interviewees in Group C-3 made a set of statements on big-data-facilitated shifts and transformations of the scope of clinical expertise. To begin with, big data could be an opportunity to purposefully widen this scope and to systematically consider data from the distinctive kinds of expertise that *patients* provide. For example, the personalization of cancer treatment is often understood as being based on objective data from genomic analyses and statistical inferences about biomarkers. However, in the spirit of expanding the *variety* of data being considered, one component of

big-data-driven improvements to our understanding of diseases and the provision of care, e.g., targeted cancer therapy, is the inclusion of more subjective, qualitative data about the patient experiences, covering their perceptions and subjective level of well-being, and thereby adding distinctive layers of evidence in addition to vitals and biomarkers by "let[ting] patients constantly reflect how well they are feeling under a therapy" [C28].

The scope of expertise of *clinicians* is affected as well. According to one line of thought, data-intensive systems could under the right conditions compensate for variances or gaps in expertise. Decision support systems that fulfill functions such as safety checks, highlighting of abnormal health parameters, or the suggestion of certain treatments are helpful especially for aspiring colleagues, including but not limited to instances where they encounter patient histories and presentations that are non-standard and in which the best course of action is difficult to determine. One of the interviewees illustrated this by means of a recent case involving a breast cancer patient in ongoing chemotherapy: in this patient encounter, it was not immediately obvious for the interviewee why a particular chemotherapy treatment had been chosen for the patient. After conversing with a number of colleagues, it turned out that the patient had an unusual combination of characteristics that warranted the treatment, and that one senior colleague could quickly point to a rationale for it. The same assessment would have been more time-consuming for less experienced clinicians. Automated systems could in principle be helpful to clinicians dealing with such non-standard cases more efficiently [C18] and leverage expert knowledge to support junior colleagues [C17]. Seemingly in contrast to this optimistic suggestion, another line of thought pointed to disadvantages: for aspiring clinicians, reliance on data-driven methods in clinical routine could actually slow down rather than enhance learning processes if these tools sometimes function as surrogates for critical reflection and independent inquiry [C23].

As a further dynamic besides support and overreliance, effects on the authority of the clinicians are conceivable if big data suggest certain diagnoses, predictions, or treatment options. One view was that big data methods "are decision aids, just like guidelines. It is exactly the same. I see it as a recommendation of a treatment which I can follow or not" [C24]. According to a different take, "for the clinicians it will become increasingly difficult to take their own decisions and to sometimes be pragmatic or to treat the individual patient since of course everything one does is constantly compared to the standard" [C19]. Reasons were also given why clinicians retain final responsibility even if datafication and automation suggest certain treatments. One of these reasons relates to a principled epistemic challenge of working with general, statistical, and/or population-level data: even the largest amount of available data of such kinds inevitably leaves certain gaps as they do not indicate or determine how the individual patient relates to and can be subsumed under such data. Precisely because it is the clinician's judgement that subsumes the particular case under the generality of the evidence base, they will be—and have to be—answerable for outcomes [C19, C29]; "as always, in the end the clinician is liable. This is just how it is. […] the patient takes precedence and if anything is different there, which forces me to take a different decision, then I do have to take this into account" [C24]. This leads a second, related reason for ascribing responsibility for big-data-informed decision-making to clinicians: suggestions provided by the decision aid will—and ought not—be followed automatically. In cases of doubt about the accuracy of the output, consultation and joint deliberation with more senior colleagues is necessary, e.g., to retrace the rationale behind the output, to enable the clinician to consider and assess the suggestion, and to arrive at a balanced decision about the best course of action [C31].

4 Discussion

Our study does have limitations. The views were gathered as reactions to the inputs, cases, and interview guides outlined in the *Methods*, and care should be taken when extrapolating or generalizing from these reactions. Moreover, clustering should not be taken to suggest statistically significant differences in attitudes amongst the groups, just that in the context of our sample, these argumentative patterns occurred in this group. No claims about uniqueness or exhaustiveness can be made on this basis. Moreover, both patients and clinicians in this study are situated in a healthcare institution with a specific focus (a Women's Hospital), within a department focusing on cancer treatments. This specific profile raises further questions, beyond the scope of the present manuscript, on the extent to which other clinician and patient groups share similar perspectives. As one exemplary hypothesis, the idea that big data and AI can lend "hope" to present and future patients might be particularly salient to patients undergoing or facing cancer treatments. Our manuscript makes a contribution to the evidence base for such claims, but does not by itself support such generalized inferences. Further empirical investigations are needed to arrive at such cross-contextual analyses.

As for the methods used in this study, the necessary transition to telephone interviews doubtlessly posed challenges. Guides that were developed for in-person group discussions, e.g., the inputs on Cases 1–3, had to be adapted to one-on-one phone conversation. Rapport is sometimes not built as seamlessly, e.g., due to a lack of visual cues. Potentially as a result, some interviewees were initially less forthcoming. Despite these limitations, the decision to transit still seemed warranted in view of the non-ideal circumstances

during the time of the study. We were encouraged by researchers defending the utility and potentials of qualitative telephone interviews (Novick 2008; Drabble et al. 2016), e.g., decreased intrusiveness, which seemed relevant to the patient population from which the sample was drawn.

While the academic discourse (as sketched in the *Introduction*) on clinical big data and AI understandably focuses on cutting-edge technologies that push boundaries and promise groundbreaking benefits, some of our results suggest that notwithstanding the potentials of big data in the clinic, in many contexts it would already be beneficial to get *small* data right. High-fidelity, personalized, and/or automated data-driven tools, e.g., in clinical decision support, are the subject of clinical research but remain largely hypothetical in routine care. Indeed, parts of the result suggest that the hype around clinical big data is not (yet) matched consistently by its impacts on the ground. Setting aside research activities, currently most relevant to improvements of the patient experience and clinical workflows in the study site and the health system in which it is embedded are issues surrounding the digitization of patient records and their trusted exchange in and between healthcare sites. Clinicians consistently mentioned the desirability of EHR to supersede cumbersome paper-based records. In the patient experiences as well, the issues of efficient sharing of information within and between hospitals as well as avoiding redundant diagnostic measures were raised.

Even for the cutting-edge applications, one constructive reminder from the clinician interviews is that at least some of their putatively transformative features are continuous with established tools. The Oxford CTG, a widely used application that has been in development since the late 1970s, was mentioned as an example for a clinical big data application already in use. It does share features with more cutting-edge applications, e.g., real-time, automated readout of parameters and assessments based on comparisons with a large reference database. Consideration of existing practices around such established data-intensive technologies could guide approaches to new technologies as well as the specification of their distinctive, novel features.

As for an apparent similarity between established and novel data-driven clinical processes and technologies, both the patient and clinician perspectives overlap and cohere with regard to expectations about the role of clinicians and the significance of human judgement when using such tools: whether it is the Oxford CTG, a digital tumor board, or a gene expression test, even with the highest data intensity, the generated outputs are not authoritative or meaningful in themselves, but precede interpretative processes driven by human decision-makers (resonating with, e.g., Lupton 2013, van der Wilt et al. 2015). First, "bare figures" need to be interpreted and contextualized to become meaningful. Second, given such interpretation and contextualization, particular courses of action need to be weighed and chosen. These findings about the perceived need for, firstly, the interpretation of data and, secondly, the forming of decision-oriented judgement in view of such interpretation coheres with results from other empirical studies of stakeholder experiences with digital health. These studies suggest that in view of increasing data intensity, clinicians proclaim to take on roles vis-à-vis the patient and the technology of interpreting, contextualizing, supervising, and counselling in shared decision-making (e.g., Jongsma et al. 2021).

A prominent concern about big data and AI in health is bias. Bias can come in various forms (Danks and London 2017). For example, a tool can be biased in the sense of somewhat systematically failing to capture or to align with reality. Bias in this sense was mentioned by the interviewees, e.g., in patients' skepticism about the reliability of Google Flu Trends. A more specific sense of bias in the sense of implicitly or explicitly discriminating against some groups receives wide attention in the academic literature on big data and AI and health (e.g., Obermeyer et al. 2019). Interestingly, interviewees did not mention this as a concern, suggesting that the importance of this issue for academics reflecting on macro- and meso-effects of such technologies need not translate to clinical stakeholders at the micro-level.

Departing from this common ground in interviewee perceptions, two areas can be highlighted in which established practices and schemes appear to reach their limits and big data and AI in health unfold disruptive potential. A first domain of interest is the status of human decision-makers vis-à-vis machine-driven outputs. Patients expect clinicians to interpret and to counsel, while clinicians appreciate data-intensive tools functioning as safety nets, highlighting potentially unconsidered treatment options, and representing knowledge akin to medical guidelines and experienced experts. Yet, clinicians emphasize that it is them who bear the final responsibility for outcomes, regardless of which additional tools, mechanisms, and means of datafication are in place—a position that this also echoed in the literature (Neri et al. 2020). This combination of views raises several further questions, problematizing the intuition that it is always the clinician who bears responsibility.

To begin with, there might be scenarios in which the sophistication of a highly data-intensive and automated tool crowds out the import of the human decision-maker, who would be ill-advised, reckless, and/or blameworthy for deviating from courses of action that in view of the respective, sophisticated (ex hypothesis) output appear advisable. Epistemically, such a scenario would at the very least pose a burden of proof for the deviating clinician (Braun et al. 2020; Kempt et al. 2022). For favorable health outcomes facilitated by the tool, there is a question to what extent we can consider the clinician truly responsible for these outcomes.

As an extension of the very same scenario, suppose the clinician does carry out the action that in view of the output from the sophisticated tool appears advisable, but an untoward health outcome results. In health in particular, ascribing and taking responsibility for error serves important purposes related to consolation, solace, understanding, hope for improvements, respect, and trust in the patient-clinician relationship (Tigard 2019). The importance of these purposes might warrant upholding responsibility, blame, and answerability even in cases and situations in which addressees could in principle invoke excuses (ibid.), e.g., being overburdened, operating with powerful yet somewhat opaque technology (Amann et al. 2022), or pointing to the best evidence and best practices that were being followed. Still, a residual challenge remains with holding the clinician responsible for following the course of action that seems epistemically ideal in the particular context under consideration.

This possibility raises the question who else besides the clinician could bear responsibility for health outcomes like those in the scenario. Pending an answer, responsibility gaps loom, i.e., situations and/or aspects for which no human agent can plausibly be deemed morally or legally responsible (Matthias 2004). Naturally, the focus would shift toward human agents in the wider ecosystem, such as developers, regulators, and leadership at and in between the hospital and health system levels. The details of widening the scope of addressees of responsibility ascriptions continues to be the focus of attention (Nissenbaum 1996; Poel et al. 2015). If a full-blown responsibility gap is avoided, still partial shifts and mismatches expectations, ascriptions, and recognition of responsibility in clinical contexts are conceivable (Bleher and Braun 2022). Considering (solely) the clinician responsible for the untoward health outcome threatens to be an oversimplification of an organizationally and epistemically complex interplay between various human decision-makers and the clinical technology under consideration.

A second area in which health-related big data and AI challenge established practices and schemes is the specification of conditions under which it is admissible to draw health-related inferences. The patient reactions to Google Flu Trends and mental health assessments based on Instagram data display two themes. First, a perceived unavoidability of health-related inferences, especially on the basis of data that do not initially seem to have health-related import. Second, lack of transparency about when, by whom, and on what grounds, and for which purposes such inferences can be drawn. As for the desirability of health-related inferences, the purpose of data processing is deemed particularly relevant: accessibility of data and the drawing of health-related inferences are to be welcomed if and to the extent to which they are put to effective use, benefit others by curing disease and/or saving lives (Prainsack and Buyx 2012; Hummel and Braun 2020; Braun and Hummel 2022), and interestingly do not generate disproportionate profit for private industry processors. This being said, patient reactions exhibit a considerable degree of pragmatism in the sense that imperfect information is tolerated to the extent that there is sufficient reason for confidence in the satisfaction of these desiderata. There was variance in concrete demands that followed from this, ranging from insistence on the need for consent requirements under all circumstances to a willingness to have 'their' data being accessed without consent if the conditions are appropriate, mirroring bioethical debates on appropriate modes of consent for the processing of biomedical big data (Richter et al. 2019; Ploug and Holm 2016).

Once again, there is coherence with findings from other empirical work. For example, in their review of empirical studies on patient and public views of sharing health data for research, Kalkman et al. conclude that "a social license for data-intensive health research cannot simply be presumed" (Kalkman et al. 2019), i.e., that it is not a given that patients and the public are in favor of the sharing of their data and related research. Instead, such approval turns on a variety of conditions related to privacy, risk minimization, data security, transparency, control, information, trust, responsibility, and accountability (ibid.).

Of particular note in the present study is the asymmetry in how patients perceive the cases to satisfy these criteria: mostly critical evaluations of the Google- and Instagram-based applications on the one hand, and the much more favorable judgements on data-intensive cancer research on the other. While the study was designed before the COVID-19 pandemic, the intuitions on Google Flu Trends clearly resonate with debates on COVID-19 tracing apps in which key concerns include the proportionality of benefits to private-sector processors (Sharon 2021), privacy concerns (Altmann et al. 2020), and beliefs about the app's effectiveness (Kozyreva et al. 2021). Likewise, concerns about mental health assessments on the basis of Instagram data resonate with calls from ethicists and legal scholars for disclosure, consent mechanisms, and ethics review in connection with suicide risk detection systems on social media platforms (Celedonia et al. 2021). Much less prominent in the perspective of interviewees was the potential social value generated on this basis. For example, if implemented and monitored carefully, applications of digital epidemiology, whether at individual or more granular levels, could help "to mitigate and prevent disease, and to promote public health" (Salathé 2018). Likewise, if implemented carefully, social-media-based tools could in principle enhance the accessibility of mental healthcare (Nilsen et al. 2022). At the same time, the social value was deemed to be more obvious in the case of personalized cancer research, apparently with little awareness that individualized therapy trials can fail to yield generalizable knowledge (Kane et al. 2021), or that in personalized medicine the issue of "just profits" (Prainsack

2017) remains live. This suggests a risk that the utility of internet- and social-media-based health-related inferences could be underestimated, while comparably high levels of confidence in the cancer use case could be disappointed in view of de facto challenges in aligning with the outlined stakeholder expectations.

The present study cannot conclusively determine what gives rise to this asymmetry. It is possible that some of the discussed tools for drawing health-related inferences are still too remote to the awareness and lived experiences of lay users. Such remoteness can invite hesitation and doubt. As for measures that can be recommended in light of the asymmetry, the results once again underscore the importance of proactive communication and engagement with stakeholder expectations. Moreover, many of the strategies mentioned by interviewees are systemic, and not just operating at the individual hospital levels or physician–patient encounters. For example, the introduction of EHR requires coordination at the systemic level, not only at the level of individual hospitals adapting their own IT infrastructures. Likewise, reshaping (if necessary) legal and moral responsibility ascriptions for clinical outcomes with big-data-driven tools will not be the feat of individual stakeholders. Both points suggest that system-wide, unified approaches on issues ranging from IT standards (Lehne et al. 2019) to continuous societal debate, systemic oversight, and governance at the conceptual, regulatory, and policy levels (Vayena and Blasimme 2018).

5 Conclusion

The effects of increasing datafication and automation of health-related applications continue to affect the views and interests of various stakeholders. In this study, we gather and compare attitudes and perceptions of patients and clinicians on emerging opportunities for drawing health-related inferences. Both patients and clinicians indicate that the consistent rollout and networking of EHR would generate considerable benefits in day-to-day care. Interviewees do not categorically disapprove of methods that process not only clinical data, but also other data such as social media data as the basis for health-related inferences. However, interviewees generally expect and demand that the purposes of data processing accord with patient preferences, and that data is put to effective use to generate social value. Patients tend to be concerned about disproportionate benefits to private-sector organizations. Proliferation of data with import for health-related inferences and the advent of automated clinical tools shape and partially reconfigure the role of clinicians, who take on ascribed and self-proclaimed responsibilities to guide patients in the collaborative consideration of this enriched information base. Interviewees emphasize that the treating clinicians maintain ultimate responsibility for clinical outcomes that result from decisions informed by health-related big data, an intuition that—as we have suggested—could eventually be disrupted by the putative ever-increasing sophistication of data-intensive and AI-driven clinical tools. We suggest that the tendency to maintain individualistic ascriptions of responsibility for clinical outcomes might have to be complemented with systemic governance to promote alignment with stakeholder expectations in data-intensive health settings and other domains informing health-related inferences.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00146-023-01825-8.

Acknowledgements The funders had no influence on the study's design, analysis, and evaluation.

Author contributions PH, MB, and PD conceptualized the study. PH and MB led the planning, execution, and data analysis. PH carried out the interviews and focus group discussions and drafted the manuscript, incorporating continuous feedback from MB. SB and DS contributed to the data analysis as well as the conceptualization and the writing of the manuscript. PF and KS provided feedback on the study design, facilitated the empirical investigations in which the data was gathered, and provided feedback on versions of the manuscript. All authors approved the final version of the manuscript.

Funding This work was supported by grants from the German Federal Ministry of Health (ZMV/1–2517 FSB 013). Parts of P.H.'s work took place within the research program Ethics of Socially Disruptive Technologies, which is funded through the Gravitation program of the Dutch Ministry of Education, Culture, and Science and the Netherlands Organisation for Scientific Research (NWO grant number 024.004.031). The funders had no influence on the study's design, analysis, and evaluation.

Data availability statement A subset of the data generated during this study is provided as a supplementary file to this article. Further data generated and analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest Matthias Braun declares that he received funding in the past 36 months from the German Research Foundation, the German Ministry of Health, and the German Ministry of Research and Education. Peter A. Fasching reports personal fees from Novartis, grants from Biontech, grants and personal fees from Pfizer, personal fees from Daiichi-Sankyo, personal fees from Astra Zeneca, personal fees from Eisai, personal fees from Merck Sharp & Dohme, grants from Cepheid, personal fees from Lilly, personal fees from Pierre Fabre, personal fees from SeaGen, personal fees from Roche, personal fees from Agendia, personal fees from Sanofi Aventis, and personal fees from Gilead. Patrik Hummel, Serena Bischoff, David Samhammer, Katharina Seitz and Peter Dabrock declare no conflicts of interest.

Ethical approval This study received ethics approval from the Ethics Commission of the Medical Faculty of Friedrich-Alexander-Universität Erlangen-Nürnberg (project number 453_18 B).

References

Altmann S, Milsom L, Zillessen H, Blasone R, Gerdon F, Bach R et al (2020) Acceptability of app-based contact tracing for COVID-19: cross-country survey study. JMIR Mhealth Uhealth 8:e19857. https://doi.org/10.2196/19857

Amann J, Vetter D, Blomberg SN, Christensen HC, Coffee M, Gerke S et al (2022) To explain or not to explain?—artificial intelligence explainability in clinical decision support systems. PLOS Digital Health 1:e0000016. https://doi.org/10.1371/journal.pdig.0000016

Andreu-Perez J, Poon CCY, Merrifield RD, Wong STC, Yang G-Z (2015) Big data for health. IEEE J Biomed Health Inform 19:1193–1208. https://doi.org/10.1109/JBHI.2015.2450362

Baškarada S (2014) Qualitative case study guidelines. TQR. https://doi.org/10.46743/2160-3715/2014.1008

Berthelsen CB, Grimshaw-Aagaard S, Hansen C (2018) Developing a guideline for reporting and evaluating grounded theory research studies (GUREGT). Int J Health Sci 6:64–76

Bleher H, Braun M (2022) Diffused responsibility: attributions of responsibility in the use of AI-driven clinical decision support systems. AI Ethics. https://doi.org/10.1007/s43681-022-00135-x

Böhm A (2019) Theoretisches Codieren: Textanalyse in der Grounded Theory. In: Flick U, von Kardorff E, Steinke I, editors. Qualitative Forschung. Ein Handbuch, Reinbek: Rowohlt

Braun M, Hummel P (2022) Data justice and data solidarity. Patterns 3:100427. https://doi.org/10.1016/j.patter.2021.100427

Braun M, Hummel P, Beck S, Dabrock P (2020) Primer on an ethics of AI-based decision support systems in the clinic. J Med Ethics. https://doi.org/10.1136/medethics-2019-105860

Celedonia KL, Corrales Compagnucci M, Minssen T, Lowery WM (2021) Legal, ethical, and wider implications of suicide risk detection systems in social media platforms. J Law Biosci. https://doi.org/10.1093/jlb/lsab021

Chandra Y, Shang L (2019) Inductive coding. In: Chandra Y, Shang L (eds) Qualitative research using R: a systematic approach. Springer Nature Singapore, Singapore, pp 91–106

Chen JH, Asch SM (2017) Machine learning and prediction in medicine—beyond the peak of inflated expectations. N Engl J Med 376:2507–2509. https://doi.org/10.1056/NEJMp1702071

Chin-Yee B, Upshur R (2019) Three problems with big data and artificial intelligence in medicine. Perspect Biol Med 62:237–256. https://doi.org/10.1353/pbm.2019.0012

Coiera E (2018) The fate of medicine in the time of AI. The Lancet 392:2331–2332. https://doi.org/10.1016/S0140-6736(18)31925-1

Danks D, London AJ (2017) Algorithmic Bias in Autonomous Systems. Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, Melbourne, Australia: International Joint Conferences on Artificial Intelligence Organization. pp. 4691–7. https://doi.org/10.24963/ijcai.2017/654

Darcy AM, Louie AK, Roberts LW (2016) Machine learning and the profession of medicine. JAMA 315:551–552. https://doi.org/10.1001/jama.2015.18421

Dawes GS, Moulden M, Redman CWG (1996) Improvements in computerized fetal heart rate analysis antepartum. J Perinat Med 24:25–36. https://doi.org/10.1515/jpme.1996.24.1.25

Diefenbach T (2008) Are case studies more than sophisticated storytelling?: methodological problems of qualitative empirical research mainly based on semi-structured interviews. Qual Quant 43:875. https://doi.org/10.1007/s11135-008-9164-0

Drabble L, Trocki KF, Salcedo B, Walker PC, Korcha RA (2016) Conducting qualitative interviews by telephone: Lessons learned from a study of alcohol use among sexual minority and heterosexual women. Qual Soc Work 15:118–133. https://doi.org/10.1177/1473325015585613

Emanuel EJ, Wachter RM (2019) Artificial intelligence in health care: will the value match the hype? JAMA 321:2281–2282. https://doi.org/10.1001/jama.2019.4914

Fernau S, Schleidgen S, Schickhardt C, Oßa A-K, Winkler EC (2018) Zur Rolle und Verantwortung von Ärzten und Forschern in systemmedizinischen Kontexten: Ergebnisse einer qualitativen Interviewstudie. Ethik Med 30:307–324. https://doi.org/10.1007/s00481-018-0494-8

Fiske A, Buyx A, Prainsack B (2020) The double-edged sword of digital self-care: physician perspectives from Northern Germany. Soc Sci Med 260:113174. https://doi.org/10.1016/j.socscimed.2020.113174

Gabriels K, Moerenhout T (2018) Exploring entertainment medicine and professionalization of self-care: interview study among doctors on the potential effects of digital self-tracking. J Med Internet Res 20:e10. https://doi.org/10.2196/jmir.8040

German Ethics Council (2017) Big Data and Health. Data Sovereignty as the Shaping of Informational Freedom (Executive Summary & Recommendations). Berlin: German Ethics Council

Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L (2009) Detecting influenza epidemics using search engine query data. Nature 457:1012–1014. https://doi.org/10.1038/nature07634

Glaser BG, Strauss AL (1967) The discovery of grounded theory. Strategies for qualitative research. New York: Aldine de Gruyter

Goldhahn J, Rampton V, Spinas GA (2018) Could artificial intelligence make doctors obsolete? BMJ. https://doi.org/10.1136/bmj.k4563

Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A et al (2016) Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 316:2402–2410. https://doi.org/10.1001/jama.2016.17216

Hummel P, Braun M (2020) Just data? Solidarity and justice in data-driven medicine. Life Sci Soc Policy 16:8. https://doi.org/10.1186/s40504-020-00101-7

Hummel P, Braun M, Augsberg S, Dabrock P (2018) Sovereignty and data sharing. ITU J ICT Discov. https://doi.org/10.1177/2053951720982012

Hummel P, Braun M, Augsberg S, von Ulmenstein U, Dabrock P (2021) Datensouveränität. Governance-Ansätze für den Gesundheitsbereich. Springer VS

Jongsma KR, Bekker MN, Haitjema S, Bredenoord AL (2021) How digital health affects the patient-physician relationship: an empirical-ethics study into the perspectives and experiences in obstetric care. Pregnancy Hypertens 25:81–86. https://doi.org/10.1016/j.preghy.2021.05.017

Kalkman S, van Delden J, Banerjee A, Tyl B, Mostert M, van Thiel G (2019) Patients' and public views and attitudes towards the sharing of health data for research: a narrative review of the

empirical evidence. J Med Ethics. https://doi.org/10.1136/medethics-2019-105651
Kane PB, Bittlinger M, Kimmelman J (2021) Individualized therapy trials: navigating patient care, research goals and ethics. Nat Med 27:1679–1686. https://doi.org/10.1038/s41591-021-01519-y
Kempt H, Heilinger J-C, Nagel SK (2022) "I'm afraid I can't let you do that, Doctor": meaningful disagreements with AI in medical contexts. AI & Soc. https://doi.org/10.1007/s00146-022-01418-x
Köngeter A, Schickhardt C, Jungkunz M, Bergbold S, Mehlis K, Winkler EC (2022) Patients' willingness to provide their clinical data for research purposes and acceptance of different consent models: findings from a representative survey of patients with cancer. J Med Internet Res 24:e37665. https://doi.org/10.2196/37665
Kozyreva A, Lorenz-Spreen P, Lewandowsky S, Garrett PM, Herzog SM, Pachur T et al (2021) Psychological factors shaping public responses to COVID-19 digital contact tracing technologies in Germany. Sci Rep 11:18716. https://doi.org/10.1038/s41598-021-98249-5
Kruse J (2015) Qualitative Interviewforschung: ein integrativer Ansatz. 2., überarbeitete und ergänzte Auflage. Weinheim Basel: Beltz Juventa
La Thangue NB, Kerr DJ (2011) Predictive biomarkers: a paradigm shift towards personalized cancer medicine. Nat Rev Clin Oncol 8:587–596. https://doi.org/10.1038/nrclinonc.2011.121
Lehne M, Sass J, Essenwanger A, Schepers J, Thun S (2019) Why digital medicine depends on interoperability. Npj Digit Med 2:1–5. https://doi.org/10.1038/s41746-019-0158-1
Liu X, Rivera SC, Moher D, Calvert MJ, Denniston AK (2020) Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. BMJ. https://doi.org/10.1136/bmj.m3164
Lupton D (2013) Quantifying the body: monitoring and measuring health in the age of mHealth technologies. Crit Public Health 23:393–403. https://doi.org/10.1080/09581596.2013.794931
Martani A, Geneviève LD, Elger B, Wangmo T (2021) 2021 "It's not something you can take in your hands" Swiss experts' perspectives on health data ownership: an interview-based study. BMJ Open 11:e045717. https://doi.org/10.1136/bmjopen-2020-045717
Martinho A, Kroesen M, Chorus C (2021) A healthy debate: Exploring the views of medical doctors on the ethics of artificial intelligence. Artif Intell Med 121:102190. https://doi.org/10.1016/j.artmed.2021.102190
Matthias A (2004) The responsibility gap: ascribing responsibility for the actions of learning automata. Ethics Inf Technol 6:175–183. https://doi.org/10.1007/s10676-004-3422-1
Maughan T (2017) The promise and the hype of 'personalised medicine.' New Bioethics 23:13–20. https://doi.org/10.1080/20502877.2017.1314886
Mayring P (2014) Qualitative content analysis: theoretical foundation,basic procedures and software solution
Mayring P (2019) Qualitative inhaltsanalyse. In: Flick U, von Kardorff E, Steinke I (eds) Qualitative Forschung Ein Handbuch Reinbek: Rowohlt. Springer Fachmedien Wiesbaden, Wiesbaden, pp 1–17
McCoy MS, Joffe S, Emanuel EJ (2020) Sharing patient data without exploiting patients. JAMA 323:505–506. https://doi.org/10.1001/jama.2019.22354
Meuser M, Nagel U (2009) The expert interview and changes in knowledge production. In: Bogner A, Littig B, Menz W (eds) Interviewing experts. Palgrave Macmillan, Basingstoke, pp 17–42
Nagendran M, Chen Y, Lovejoy CA, Gordon AC, Komorowski M, Harvey H et al (2020) Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. BMJ. https://doi.org/10.1136/bmj.m689
Neri E, Coppola F, Miele V, Bibbolino C, Grassi R (2020) Artificial intelligence: Who is responsible for the diagnosis? Radiol Med 125:517–521. https://doi.org/10.1007/s11547-020-01135-9
Nilsen P, Svedberg P, Nygren J, Frideros M, Johansson J, Schueller S (2022) Accelerating the impact of artificial intelligence in mental healthcare through implementation science. Implement Res Pract 3:26334895221112030. https://doi.org/10.1177/26334895221112033
Nissenbaum H (1996) Accountability in a computerized society. Sci Eng Ethics 2:25–42. https://doi.org/10.1007/BF02639315
Novick G (2008) Is there a bias against telephone interviews in qualitative research? Res Nurs Health 31:391–398. https://doi.org/10.1002/nur.20259
Obermeyer Z, Powers B, Vogeli C, Mullainathan S (2019) Dissecting racial bias in an algorithm used to manage the health of populations. Science 366:447–453. https://doi.org/10.1126/science.aax2342
Pardey J, Moulden M, Redman CWG (2002) A computer system for the numerical analysis of nonstress tests. Am J Obstet Gynecol 186:1095–1103. https://doi.org/10.1067/mob.2002.122447
Ploug T, Holm S (2016) Meta consent—a flexible solution to the problem of secondary use of health data. Bioethics 30:721–732. https://doi.org/10.1111/bioe.12286
Prainsack B (2017) Personalized medicine: empowered patients in the 21st century. New York University Press, New York
Prainsack B, Buyx A (2012) Solidarity in contemporary bioethics—towards a new approach. Bioethics 26:343–350
Price WN, Cohen IG (2019) Privacy in the age of medical big data. Nat Med 25:37–43. https://doi.org/10.1038/s41591-018-0272-7
Reece AG, Danforth CM (2017) Instagram photos reveal predictive markers of depression. EPJ Data Sci 6:15. https://doi.org/10.1140/epjds/s13688-017-0110-z
Richardson JP, Smith C, Curtis S, Watson S, Zhu X, Barry B et al (2021) Patient apprehensions about the use of artificial intelligence in healthcare. Npj Digit Med 4:1–6. https://doi.org/10.1038/s41746-021-00509-1
Richter G, Borzikowsky C, Lieb W, Schreiber S, Krawczak M, Buyx A (2019) Patient views on research use of clinical data without consent: Legal, but also acceptable? Eur J Hum Genet 27:841–847. https://doi.org/10.1038/s41431-019-0340-6
Rivera SC, Liu X, Chan A-W, Denniston AK, Calvert MJ (2020) Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. BMJ. https://doi.org/10.1136/bmj.m3210
Salathé M (2018) Digital epidemiology: what is it, and where is it going? Life Sci Soc Policy 14:1. https://doi.org/10.1186/s40504-017-0065-7
Sand M, Durán JM, Jongsma KR (2022) Responsibility beyond design: physicians' requirements for ethical medical AI. Bioethics 36:162–169. https://doi.org/10.1111/bioe.12887
Sharon T (2018) When digital health meets digital capitalism, how many common goods are at stake? Big Data Soc 5:2053951718819032. https://doi.org/10.1177/2053951718819032
Sharon T (2021) Blind-sided by privacy? Digital contact tracing, the Apple/Google API and big tech's newfound role as global health policy makers. Ethics Inf Technol 23:45–57. https://doi.org/10.1007/s10676-020-09547-x
Shaw J, Rudzicz F, Jamieson T, Goldfarb A (2019) Artificial intelligence and the implementation challenge. J Med Internet Res 21:e13659. https://doi.org/10.2196/13659
Stegemann L, Gersch M (2021) The emergence and dynamics of electronic health records—a longitudinal case analysis of multi-sided platforms from an interoperability perspective. hawaii international conference on system sciences (HICSS-54)

Suwinski P, Ong C, Ling MHT, Poh YM, Khan AM, Ong HS (2019) Advancing personalized medicine through the application of whole exome sequencing and big data analytics. Front Genet. https://doi.org/10.3389/fgene.2019.00049

Thomas DR (2006) A general inductive approach for analyzing qualitative evaluation data. Am J Eval 27:237–246. https://doi.org/10.1177/1098214005283748

Tigard DW (2019) Taking the blame: appropriate responses to medical error. J Med Ethics 45:101–105. https://doi.org/10.1136/medethics-2017-104687

Topol EJ (2019) High-performance medicine: the convergence of human and artificial intelligence. Nat Med 25:44–56. https://doi.org/10.1038/s41591-018-0300-7

van de Poel I, Royakkers L, Zwart SD (eds) (2015) Moral responsibility and the problem of many hands. Routledge, New York

van der Wilt GJ, Reuzel R, Grin J (2015) Design for values in healthcare technology. In: van den Hoven J, Vermaas PE, van de Poel I (eds) Handbook of ethics, values, and technological design. Springer, Netherlands, Dordrecht, pp 717–738. https://doi.org/10.1007/978-94-007-6970-0_36

van Veer LJ, Bernards R (2008) Enabling personalized cancer medicine through analysis of gene-expression patterns. Nature 452:564–570. https://doi.org/10.1038/nature06915

Vayena E, Blasimme A (2018) Health research with big data: time for systemic oversight. J Law Med Ethics 46:119–129. https://doi.org/10.1177/1073110518766026

Wallden B, Storhoff J, Nielsen T, Dowidar N, Schaper C, Ferree S et al (2015) Development and verification of the PAM50-based Prosigna breast cancer gene signature assay. BMC Med Genom 8:54. https://doi.org/10.1186/s12920-015-0129-6

Wilkinson J, Arnold KF, Murray EJ, van Smeden M, Carr K, Sippy R et al (2020) Time to reality check the promises of machine learning-powered precision medicine. Lancet Digital Health. https://doi.org/10.1016/S2589-7500(20)30200-4

Yin RK (2018) Case study research and applications: design and methods, 6th edn. SAGE, Los Angeles

Yu K-H, Beam AL, Kohane IS (2018) Artificial intelligence in healthcare. Nat Biomed Eng 2:719–731. https://doi.org/10.1038/s41551-018-0305-z