DISS. ETH NO. 27589

Ethics and Governance of Big Data in Health Research and Digital Health Applications

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH
(Dr. sc. ETH Zurich)

presented by

Agata Ferretti

M.A., Università degli Studi di Milano
M.Sc., London School of Economics and Political Science

born on 10.07.1992
citizen of Italy

accepted on the recommendation of

Prof. Dr. Effy Vayena, examiner
Dr. Marcello Ienca, co-examiner
Prof. Dr. Samia Hurst, co-examiner

2021

Acknowledgments

This thesis is the cumulative result of a research journey that lasted more than three years and involved the collaboration and support of many people to whom I am deeply grateful.

First and foremost, I would like to express my deepest gratitude to my supervisor, Prof. Effy Vayena, whose encouragement, expert guidance and trust in my abilities have been essential to accomplish this work. Effy, thank you for all the precious advice over the years, and for involving me in so many exciting projects. To my co-examiner Dr. Marcello Ienca, an extraordinary mentor and collaborator, thank you. I am grateful to you for accompanying me along this journey, unreservedly sharing your knowledge, and challenging my thinking to improve the quality of my research. I am also very thankful to Prof. Samia Hurst, who kindly agreed to be my second co-examiner, for her constructive feedback and terrific collaboration on the Swiss National Science Foundation (SNSF) project "BEHALF". I am very appreciative of the doctoral studies support spearheaded by the SNSF.

I am extremely grateful to all members – former and current – of the Health Ethics and Policy Lab for the countless inspiring conversations and for creating such a positive work environment. Huge thanks go to my closest collaborators, Alessandro, Caroline, Felix, James, Joanna, Julia, and Manuel for supporting me on those days when my confidence wavered. Your unfailing and valuable advice helped me to overcome the challenges and succeed in my work. I am also very thankful to all the international colleagues who contributed to specific sections of this dissertation and made my doctoral studies such a formative and enriching experience.

On a personal note, I am terribly grateful to my parents for their unconditional love, constant encouragement, and for giving me the opportunity to pursue my interests. To my wonderful partner Jan, thank you for always being by my side - and not only because you were locked at home with me during the COVID-19 pandemic! Your endless support was crucial to achieving this important goal, and your contagious optimism became an essential mood booster to get through these tough times. Finally, I wish to extend thanks to all the friends who shared their curiosity towards my work. Knowing that what I do resonates with your interests gave me a greater purpose and motivated me even more in my research.

As I stand at the end of this exciting journey, I cannot keep a smile off my face. The wise words of Kurt Vonnegut come to mind. He urged people to notice when they are happy, and to acknowledge it by thinking, murmuring, (in my case) exclaiming:

"*If this isn't nice, I don't know what is!*"

Abstract

The health sector is undergoing a digital revolution. Digital technologies monitor and capture data about people's health and physiology. The widespread availability of digital technologies and powerful analytical tools - such as artificial intelligence (AI) algorithms – makes it possible to generate, store, share and analyze vast amounts of aggregated data – "big data". Big data offers exciting opportunities to improve prevention, diagnostics and therapeutics. Despite its benefits however, big data and AI-enabled technologies come with a series of technical, legal and methodological challenges which have ethical implications including privacy, equity, fairness, accountability, risk assessment, and benefit distribution. Given these novel ethical implications, the existing ethical guidance and oversight mechanisms are inadequate to effectively regulate big data use and health technology development.

The first part of this thesis aims to gain a deeper insight into the ethical implications of big data and AI-enabled technologies in the health sector – specifically in health research and digital health applications. This analysis informs the second part of the thesis, which aims to determine the current state of ethical oversight mechanisms in health research and health apps in respect to big data. This goal will be achieved by mapping gaps in the existing ethical guidelines to suggest key reforms and future recommendations.

The findings of this thesis show that the ethical discourse is biased toward privacy and technical concerns, leaving other ethical considerations unexamined. This narrow perspective is also reflected in the lack of comprehensive governance and adequate oversight to ensure an ethically aligned use of data in health research and health app development. The focus on data security compliance may result in insufficient guidance for stakeholders to make ethical choices, and thus may cause individual and collective harms. To avoid this risk, this thesis recommends updating the existing data governance and oversight mechanisms, considering a more comprehensive and robust ethical approach, which promotes shared values beyond data protection. Furthermore, this thesis includes an ethical toolkit that can guide stakeholders toward the ethically aligned use of big data. Only within good and ethical governance, big data and AI will be able to unlock their full potential for benefit to the health sector.

Sunto

Il settore sanitario sta attraversando una rivoluzione digitale. Le tecnologie digitali monitorano e raccolgono dati sulla salute e la fisiologia delle persone. L'ampia disponibilità di tecnologie digitali e di potenti strumenti analitici – come gli algoritmi di intelligenza artificiale (IA) – permettono di selezionare, memorizzare, condividere e analizzare enormi quantità di dati aggregati – i cosiddetti "big data". L'uso dei big data offre interessanti opportunità per migliorare la prevenzione, la diagnostica e la terapia. Oltre a tali i benefici, i big data e le tecnologie di IA comportano anche una serie di sfide tecniche, legali e metodologiche che hanno implicazioni etiche per concetti come privacy, giustizia, responsabilità, valutazione di rischi ed equa ripartizione degli effetti positivi. Date tali nuove implicazioni etiche, le indicazioni normative in vigore ed i meccanismi di sorveglianza sono inadeguati a regolare efficacemente l'uso dei big data e lo sviluppo delle tecnologie sanitarie.

La prima parte di questa tesi mira a sviluppare un approfondimento delle implicazioni etiche dei big data e delle tecnologie di IA nel settore sanitario - in particolar modo nella ricerca biomedica e nelle applicazioni digitali per la salute (ovvero le "app sanitarie"). Questa analisi introduce la seconda parte della tesi, il cui scopo è quello di individuare qual è lo stato attuale dei meccanismi di vigilanza etica nella ricerca biomedica e nelle app sanitarie rispetto alle novità dei big data. Questo obiettivo sarà raggiunto elencando le lacune esistenti nell'orientamento etico per poi suggerirne riforme chiave e raccomandazioni future.

I risultati individuano che l'attenzione etica è ampiamente incentrata sulla privacy e sulle preoccupazioni tecniche, lasciando inesplorate innumerevoli altre considerazioni. Questa prospettiva ristretta si concretizza anche nella mancanza di una governance esaustiva e di una supervisione etica adeguata, che possa garantire un uso corretto e rispettoso dei dati nella ricerca sanitaria e nello sviluppo di app sanitarie. Concentrarsi sulla conformità alla sicurezza dei dati potrebbe risultare insufficiente per guidare le parti interessate a fare scelte etiche opportune, e di conseguenza potrebbe comportare effetti negativi sia individuali che collettivi. Per evitare questo rischio, questa tesi raccomanda di aggiornare la governance dei dati e i meccanismi di supervisione alla luce di un approccio etico più completo e solido, che promuova valori comuni al di là della semplice protezione dei dati. Inoltre, la tesi include uno strumento etico di concreto utilizzo che nella pratica indirizzi le parti interessate verso ad un uso dei big data eticamente allineato. Solo all'interno di una più ampia ed etica governance, i big data e l'IA saranno in grado di esprimere tutto il loro potenziale e apportare benefici al settore sanitario.

Table of Contents

“Cram them full of non-combustible data, chock them so damned full of 'facts' they feel stuffed, but absolutely 'brilliant' with information. Then they’ll feel they’re thinking, they’ll get a sense of motion without moving. And they’ll be happy, because facts of that sort don’t change. Don’t give them any slippery stuff like philosophy or sociology to tie things up with. That way lies melancholy.”

Ray Bradbury, Fahrenheit 451

CHAPTER 1: General Introduction

1.1 The big data revolution in the health sector

1.1.1 *Big data and the digitalization of the human*

Everywhere we turn, digital technology is constantly there. And it never sleeps. Not even when we sleep, it stops interacting with us (1).
The digital life that people today – particularly in advanced economies (2) – take for granted was only a fantasy at the beginning of the 21st century (3). Since then, the progressive expansion of telecommunication infrastructure, as well as the World Wide Web has paved the way for an unprecedented transformation. This digital transformation has affected human activities (financial, economic, and social) and countries around the world to varying degrees(4). Nonetheless, we can observe a general growth in the availability and use of digital technologies (e.g., mobile phones and access to the Internet) on a global scale.

The growth of the global digital economy resulted in a rapid decline of human activities that could not adapt to this radical technological shift – namely those which did not accumulate data (5). The accumulation of data is the main currency of the digital economy (6), and its mantra, “Data is the new oil”, proves it. In fact, digitalized capitalism profits from the commodification of life itself (7), as it profited from oil in the twentieth century. Indeed, actors from many different sectors including business, politics, marketing, basic science, health and transportation realized that information derived from human data has an incredible predictive power, and that these predictions can be exploited for financial gain (8). In light of this, every aspect of human life has been digitalized, and digital traces are disseminated everywhere (9). These traces include data describing online interpersonal transactions and interactions, as well as data concerning individuals’ characteristics and preferences. The latter data are often created through digital infrastructures such as the Internet of Things (IoT). The minds and bodies of humans themselves have been progressively quantified by interacting with these technologies (10). The complex data are collected both directly, by individuals (e.g., stepping on a scale or doing a genetic test), and indirectly, by software running silently in the background of wearable devices and smart sensors (11). Consequently, in 2018 the European General Data Protection Regulation (GDPR) popularized the term “data subject”, meaning an individual identifiable through data points (12).

Data volume and variety has grown exponentially thanks to the widespread availability and uptake of digital technologies (13). Moreover, the pervasiveness of the Web has allowed data to be

collected and shared in unprecedented volumes, popularizing the expression *"big data"* (14, 15). The literature often refers to the three Vs (volume, variety, and velocity) as the core characteristics of big data (16), although consensus is still lacking on a clear definition (17). Some critics highlighted that an additional essential feature of the big data revolution is the availability and deployment of powerful analytics tools (18). The role played by Artificial Intelligence (AI) – defined as "the automation of intellectual tasks normally performed by humans" (19)– has become central to the big data debate. Indeed, data alone would not be so precious if it were not for the tools that can extract valuable insights from it.

AI algorithms are developed and trained using big data, to make predictions which advise the decision making of involved actors. For instance, it is common knowledge that social media uses big data and AI for targeted advertising and to influence users' shopping habits. This is one example of how the line dividing human life (in its social, economic, biological, and physical aspects) and the digital world (made up of online networks, algorithms, data, and technologies) is increasingly blurred. Klaus Schwab (the founder of the World Economic Forum) referred to this progressive and unstoppable integration of the two worlds as a "fourth industrial revolution." These two dimensions, which until recently existed alongside each other, are merging to become one (20). As a result, a complex reality arises, bringing radical change to how humans interact, produce, consume, communicate, learn, and live. Considering this, we must discuss the ethics and governance of big data in the health sector.

1.1.2 *Big data and the digitalization of health*

Digitalization is revolutionizing the health sector. For example, Electronic Health Records (EHRs) are replacing traditional handwritten medical records in many countries around the world (21, 22). This is often due to efficiency reasons: EHRs improve the process of storing and accessing copious amounts of data beyond the patients' medical records (including demographics, medical imaging, genetic test results, and family history.) (23). The widespread adoption of EHRs has made them the primary source of biomedical big data. However, the emergence of affordable genetic testing and sequencing, too, contributed to the explosion of big data in the health sector (24). Increasingly, large volumes of genetic data are collected through public programs, such as the 100.000 Genomes Project established in 2018 to sequence 100K genomes from National Health Service (NHS) patients with rare diseases and cancer (25). These programs are not

exclusive to genomics, taking place in other 'omics sectors such as proteomics, epigenomics, transcriptomics and pharmacogenomics well a, etc.(26). Meanwhile, private companies have entered the space of biomedical big data through direct-to-consumer health-related products (27, 28). A notorious and pioneering example of this is the US company 23&Me, which had collected and genotyped the DNA of more than one million customers by 2015 (29). Furthermore, commercial entities have increased production of wearable devices, ambient assisting living (AAL) tools, medical imaging technologies and other digital health technologies which collect, analyze, store, and share large volumes of data (30).

Although a precise and universally accepted definition of "digital health" does not exist, the US Food and Drug Administration (FDA) includes five categories within this concept. Among these categories we find: (i) health applications available on mobile phones, also known as *mobile health*, *mHealth*, or - more commonly - *health apps*; (ii) wearable devices and sensors ; (iii) telehealth and telemedicine, meaning remote clinical care, educational activities and public health initiatives administered through telecommunication technologies; (iv) Health Information Technologies (HIT), namely EHRs or other informatic infrastructures made to collect, protect, and retrieve administrative, financial, and biomedical data; and (v) personalized medicine, a form of medical care that uses technology to tailor treatment and therapy to best match individuals' needs. Digital health is characterized by a feedback data flow that goes from the patients/users to health care professionals and health services, and then back to the patients/users.

That said, the FDA's description of digital health is not the only one. Some authors present digital health as complementary to digital medicine (31, 32). Within this conceptualization, digital health means that individuals monitor their own health and fitness status using personal digital tools (such as smartphones or wearables with integrated sensors). Digital health applications are intended to be low-risk and developed for direct-to-consumer use. Digital medicine, however, refers to potentially higher-risk medical devices that undergo rigorous clinical tests and validations. Once approved, these devices can complement clinical practice by preventing, diagnosing, and treating disease. Many countries implement this definition to distinguish between health apps and medical apps (33). While the former are loosely assessed, as they only function as tools for monitoring and changing behavior, the latter have a specific clinical purpose and are regulated as medical devices.

This thesis will use this second definition of digital health. This definition helps to focus on the role played by smartphones and wearables, as well as by smart home devices and sensors, to monitor the health status of individuals. For example, smartwatches track users' daily fitness activity with real time data. Specifically, using biosensors, the watches record vital parameters such as heartbeat and blood oxygen level. Digital health often acquires data in a passive way (without active inputs by the user) (34). To be "invisible" to the user is particularly useful when monitoring patients affected by mental health disorders (such as depression or anxiety) (35), or cognitive decline (36). In this regard, smart homes powered by the IoT are particularly useful to monitor the health of the elderly. Sensors on refrigerator doors and kitchen stoves, in beds, and under the sofa can reveal eating habits, sleeping patterns, and mobility tendencies respectively (37).

User-related data are constantly produced in the digital society (38); these include social media content, Internet searches, online purchases, fitness class attendance, ambient noise levels, user voice metrics, technology usage patterns (such as typing speed or reaction time to visual stimuli), and location information (39). "Digital phenotype" refers to this quantification of the health status of a person using smartphones and other digital technologies (40). Today, researchers and private corporations can infer health information even from these non-biomedical data (41). For example, studies have correlated obesity with grocery data (42) and have detected suicidal intentions from Facebook posts (43). A well-known example of this is when the retail company Target – by analyzing online data – discovered that a woman was pregnant, before she had even told anyone (44). In the digital health landscape, this variety of data (which this thesis refers to as "health data") has become extremely valuable for several reasons. First, it compensates for the limited accuracy that may characterize a sole source dataset. Second, it provides a more granular description of the health conditions existing on the scale between healthy and unhealthy, as defined by the World Health Organization (45). Third, it promises to deliver better, and individually tailored healthcare (46). Finally, it helps to fill the knowledge gap on disease outbreaks and containment strategies, such as in the case of the SARS-CoV-2 pandemic (47). Regardless, health data can reach its true potential only when paired with AI analytic systems. The following subsection will explore the most promising applications of AI to improve individual health, public health, and health services.

1.1.3 Big data and AI-enabled technology: what are the opportunities?

Without AI, big data could not revolutionize the health sector. At this point in history, no human could replace AI (the converse is still open for debate) (48). The powerful role of AI is obvious. First, algorithms link data from a variety of sources; then, they extract valuable information; and finally, they draw useful conclusions based on the extracted information. Most AI-enabled technologies are currently task oriented (i.e., narrow in scope) and work as "expert systems" (49), applying learned rules to new case studies. Machine learning (ML) on the other hand, the most common form of AI, learns from the data and improves its assessments over time on its own. ML identifies patterns and correlations in the data, without necessarily offering a causal explanation (50). ML data processing offers new opportunities for understanding diseases, developing treatments, and improving health services (51).

Big data analytics has already proved its potential for diagnostics, for example, by detecting malignant moles from medical images (52). In 2017, ML successfully recognized melanoma with the same accuracy as a human dermatologist (53). Today, AI can help fight widespread diseases such as diabetes (by looking at a patient's retinal fundus) (54) and pneumonia caused by the SARS-CoV-2 virus (55). Predicting and preventing diseases another field where ML excels. ML has already been used to predict psychotic episodes from natural language data (56) and the risk of breast cancer in young women from EHR information (57). Disease prevention can save lives; the same is true for precision medicine or finding the most effective treatment for each individual patient. Precision medicine is one of the most promising fields of ML (58).Especially in oncology, AI applications allow physicians to move away from a "one size fits all" treatment approach and focus on finding the best available option for each individual patient.

In the field of digital epidemiology, AI advancements were recently put to the test. The COVID-19 pandemic pushed governments around the world to develop digital health applications to trace infected individuals and to protect public health (59). AI-powered technologies analyze data, define the health profile of an individual, and consequently inform public health authorities of disease patterns and potential outbreaks. Even before the COVID-19 pandemic, many countries around the world adopted programs of participatory disease surveillance which rely on self-reported symptoms further analyzed by AI algorithms (60). Especially in low-and-middle income countries (LMICs), the uptake of mobile phones and health apps allows researchers to collect granular, real-time, and local information. This translates into timely interventions to contain

deadly diseases, such as Ebola (61). Furthermore, public health officials increasingly rely on AI as a strategic tool to reach public health goals. For example, they can tailor public health campaigns about COVID-19 vaccination (62) or contraceptive methods (63) based on the general public's attitudes and sentiments expressed as on social media.

AI-powered health apps can increasingly assist patients in self-care practices, with a high degree of efficiency and quality (64). Consider those apps that deliver healthcare services previously administered by healthcare professionals, for example, those helping the elderly to take their prescribed medication correctly or offering psychological counselling via chatbots. Hence the popularity of these technologies: they extend access to primary care while reducing the burden on physicians and family caregivers (65). AI-enabled technologies are not only fundamentally cost-effective, but they also allow healthcare systems to more efficiently allocate available resources. Today, hospitals' resource planning relies on algorithms that optimize health management by analyzing EHRs data and identifying high-cost patients or probable hospital readmissions (66).

The *health data ecosystem* (67) has evolved over the past decade to include more data, unique data sources, and powerful analytics methods. These changes resulted in a broad array of stakeholders engaging in big data research and development alongside patients, academic researchers, healthcare providers, and healthcare institutions. Among these stakeholders are health industries (including health insurers and pharmaceutical companies), professional health organizations, and citizens organizations (i.e., citizen science, patients', and consumers' associations). Private actors from the data industry also play a powerful role in this ecosystem despite their tangential connection to the health field (68). Despite having different motivations (e.g., commercial goals vs. public interest), these actors often use the same datasets. Increasingly often, consumer technology corporations, telecommunication companies, and technology startups partner with health and academic institutions to extend the boundaries of biomedical knowledge (69). This is illustrated, for example, by the agreement between the US National Institute of Health (NIH) and Fitbit to improve personalized research (46), or the UK NHS partnering with Google's DeepMind to apply AI for preventive and diagnostic purposes (70).

The examples mentioned in this section show that big data and the AI revolution in the health sector is already taking place. Accompanying this revolution, however, is a series of concerns, which will be explored in the next section.

1.1.4 Pitfalls of big data and AI-enable technology in the health sector

As the complexity of the data ecosystem increases, so do the pitfalls associated with it (71). A first potential issue concerns the distinction between public and private data. The literature discusses misunderstandings that emerge from using data available online, for example, social media posts, as 'public' data (72). Indeed, people have certain expectations about what can be done with their data, despite the fact they are accessible online (73). These expectations might consist of having the data i) safely stored and protected from unauthorized access ; ii) shared with third parties only with previous agreement and in anonymized form; and iii) re-used exclusively for purposes closely related to those for which they were originally shared (74). As some authors pointed out (75), individuals may refrain from having their personal information created in one context (for example healthcare) and transferred to be re-used in another sphere (such as commercial advertisement). This is because the data could change in meaning when shifted out of context, and thus lead to disadvantages or harms for the individuals.

In health research, access to personal information is traditionally regulated by precise rules embedded in an informed consent. In the era of big data this approach poses some problems (76). First, obtaining consent for any use and re-use of big data is highly impracticable, if not impossible given the data volume. Particularly, tracing back individuals becomes particularly unfeasible when multiple de-identified datasets are merged (77). Second, the purpose of big data analysis is often exploratory. That is, one cannot anticipate ex-ante what the connections among the various data points will say about the subjects involved, and – most importantly – whether future analyses on the same data will reveal additional information (78).

Uncertainty is a key feature of big data research and digital health (79). For instance, incomplete or unstructured data, as well as ML explorative clustering models, can present uncertainty (80). Uncertainty introduces many issues concerning the access, usage, and sharing of the data. Despite these concerns, questions about data access (who can access the information and under which conditions), data control (who controls vs. who owns the data), data sharing (who can use the data and for what purposes) and data governance (who oversees the data uses) often remain unanswered (81). Regardless, these issues become increasingly relevant considering potential future data uses, such as relationship data between subjects emerging from genetics, environmental exposures, and health activity patterns.

As previously mentioned, AI identifies patterns in the data, but not without presenting three main challenges. The first one has to do with ML pre-processing. That is how to clean and normalize scattered data to have reliable and evidence-based outcomes (82). The digital divide between countries plays a key role in the health sector. In fact, there are elevated levels of inequalities between countries that have the resources, infrastructure, and digital skills to access and use technologies, and those that do not. This is amplified by the paucity of data describing certain subgroups, which in turn limits the representativeness of the data. The digital divide also results in low digital health interoperability within and across countries due to elevated levels of data heterogeneity (i.e., data disparate in origins, purpose, format, quality, and time of collection). The second concern is closely related to the previous one and is epistemological in nature. That is, that health knowledge increasingly arises from algorithms' data classification rather than from proven scientific causality. This issue emerges from the complexity and the lack of transparency in ML algorithms and emphasizes the challenge of assessing the quality and scientific validity of ML outcomes (83). The third issue occurs on the ontological level, and consists of the dissolution of the demarcation between sickness and health (84). This happens because AI can accurately predict the risk of developing a certain disease, defining a new intermediate grey area, for those who are not yet ill but will be.

Finally, a broader problem arises from the distribution of benefits among the various stakeholders (patients, commercial companies, public institutions, vulnerable groups, the public, etc.). Notably, these big data trends in the health sector are not sufficient descriptions of the world to produce knowledge or explain a given phenomenon, and in a world where human life is translated into strings of data, the actors who own the analytic tools to interpret the data become extremely powerful.

1.2 Rationale for this thesis

The introductory section of this thesis illustrates that a comprehensive appraisal of big data and AI-enabled technologies must review both the opportunities and pitfalls to reach beyond the hype. Despite its importance, this appraisal may be insufficient to orient stakeholders' decision making in light of the emerging ethical, legal, and social implications (ELSI) of this technology. In recent years, an increasing amount of literature has addressed these implications. Particularly, the ethical debate has been very lively. Nevertheless, the variety of considerations on the topic

appeared fragmented, rather than embedded within a rigorous ethical and normative framework. Although some authors have proposed models of ethical data uses (85-87), the current absence of a comprehensive and unitarian recommendation system calls for two endeavors:

- First, to assess whether there are any gaps in the ethical debate, and if all relevant ethical implications have been thoroughly and equally considered.
- Second, to explore the state of the ethical oversight in relation to big data and AI-technologies uses in the health sector. Namely, whether technological development is – in practice – adequately guided by ethical principles and aligned with societal values.

The next subsections provide further justification for each of these two endeavors.

1.2.1 Ethical implications

The big data and AI pitfalls mentioned earlier are of a technical, legal, and social nature. Regardless, they require careful consideration considering their ethical relevance and implications.

These ethical implications force us to reflect on what societal model we want to bring forward and whether our efforts are going in the right direction. In the health sector, multiple stakeholders promote different paths forward based on their diverging incentives, leading to many questions. Which values does society want to promote? Are modern technologies and their uses aligned with these values and public interests? Are the technological risks proportional to the expected benefits, and are they fairly distributed among the stakeholders?

The problem of dismissing such questions lies in the fact that technological development is not morally neutral, despite some inventors and engineers arguing that the effects of a technology come directly from the use that people make of it, rather than from the technology itself (88). However, as many philosophers in the second half of the twentieth century remarked, technology is not just a neutral means to an end, which can be used for either good or bad (89). Philosophers such as Heidegger observed that technology is not a descriptive tool to understand the world, but an instrument designed with the precise goal to change it (90). In other words, there is no instance in which a technology can exist without shaping the world around it. On the contrary, technology and its socio-cultural consequences “legislate the condition of human existence” (91). Digital health applications, for example, are developed and deployed within goal-oriented schemes; in practice they change the world, and in doing so they carry some values. These values, however,

are neither determined by the technology itself, nor do they follow some teleological objective, therefore, they are not deterministic. Rather, cultural, economic, and political factors infuse technologies with their own values and goals. In this way, data do not simply equal facts (92). Indeed, data are not self-explanatory; they cannot speak for themselves. Instead, data require curation, analysis, and interpretation, which are inevitably conditioned by sociocultural conditions. Because of this, when interrogated, data can offer alternative representations of the world, as values shape data practices and inferences (71).

Data and technologies are not self-reflective either, despite their power to shape societies. Society, on the contrary, has the moral obligation to act responsibly and to infuse technological development with ethical values. We all should critically examine whether the big data and AI-enabled transformations in the health sector are ethically acceptable or not. Failure to fulfill this task can result in health data misuses and corresponding harms, unreliable digital health technologies, distrust in health institutions, and more health inequality (93). Conversely, health benefits for both individuals and society will emerge from an ethically aligned used of technologies.

Therefore, we need to carefully evaluate the emerging ethical considerations. The most critically debated ones are summarized below.

The literature on big data in the health sector most often highlights *privacy and confidentiality* among all the ethical implications. Violations of data subjects' privacy are ethically problematic for two reasons. First, these violations directly infringe on a fundamental right, as they breach the individuals' private sphere. Second, privacy breaches can harm individuals and groups in an indirect way: even when the data of a particular individual or group are not disclosed, their information can be inferred using big data analytics. Revealing information could potentially lead to risk of discrimination, a factor often debated alongside considerations of *bias, fairness, and transparency* of AI algorithms. For instance, a bias in the datasets or design of a digital health application could be perpetuated in the machine outcomes. In turn, these outcomes could profoundly affect the health of individuals or population subgroups by discriminating them. Unfortunately, detecting biases sometimes requires explaining how the algorithms process the data, something that is not always possible due to the complexity of the algorithms (especially those characterizing ML). In fact, the literature often refers to the ML procedure as a "black box"

since it is not easy to penetrate the decision-making processes and the internal functioning of the machine (94).

This lack of algorithmic transparency can directly influence the *trustworthiness* of the technology itself. The end users of digital health applications (such as patients, doctors, and the public) might be skeptical to trust tools they do not understand, and consequently refrain from using the tools. Indeed, new forms of *harm* (i.e., emotional, dignitary, or financial), alongside the physical one, can emerge from health data processing. However, researchers and developers often fail to predict the occurrence of these harms - due to the very nature of data analysis. Ethical questions arise in this regard: how to set the threshold for *minimal risk* in research and in using digital health applications if identifying all potential risks is implausible? How to effectively anticipate and mitigate unforeseeable big data risks? How to adequately inform data subjects of these risks?

Ethical considerations about *self-determination and individual autonomy* emerge in relation to topics such as the return of results to the patients, or data re-use for new purposes and by new actors (e.g., health data initially used for public research and then shared with private actors for commercial purposes). Some authors also reflect on the ethical meaning of *empowerment* within the digital health transformation (95, 96), asking whether it is ethical to shift the responsibility of providing healthcare from public institutions to the patients themselves. Self-care technologies – such as health apps - put individuals in charge of their own health. It has yet to be determined whether, in practice, this opportunity is fair and improves peoples' lives. It may be that - in certain conditions – this exposes patients and/or other population subgroups (e.g., those without access to digital technologies) to harm (9).

Similar discussions reflect on whether a *just distribution of benefits* is achievable when using big data and digital health applications. One risk consists of disproportionally benefiting certain stakeholders at the expense of others. For example, certain companies offering genealogical DNA testing have unlimited access to the genetic code of each customer while "only" providing ancestry information in return. There are other equity considerations in the debate of big data and AI for global health: do AI-enabled technologies provide access to healthcare for those most in need, or do they perpetuate existing inequalities? Big data challenges call for an ethical reflection on the current data *governance and technological accountability* (97). Multiple actors (individuals, groups, corporations, and institutions) share responsibility for the data uses and their consequences. Adequate data governance should ensure that the outcomes of these data uses

are fair and lead to appropriate decision making. However, are the current oversight mechanisms of data research and initiative effective in preventing harm? How should responsibility be allocated within these networks, when fraudulent, misguided, and unethical practices take place?

1.2.2 Implications for the ethical oversight mechanisms

As technological development advances – and brings unprecedented risks – novel forms of accountability and ethical oversight will be required (98). The fact is, that the ethical norms and rules promoted by governance mechanisms mirror specific historical conditions. Take, for example, research ethics committees (RECs) in the biomedical and clinical sector. RECs – also known as Ethics Review Committees (ERCs), or Institutional Review Boards (IRBs) in the US – emerged after the Second World War. They promoted ethical and normative principles (i.e., autonomy, justice, beneficence, and non-maleficence), addressing the moral dilemmas of the time, and thus, ensured that terrible experimental trials – such as those carried out during the war - would not be repeated. Likewise, the regulatory landscape for developing, testing, and validating pharmaceuticals or medical devices emerged in a pre-big data context and prior to the widespread diffusion of AI-enabled technologies (99). Therefore, the existing oversight mechanisms might be unfit for ethically reviewing new research and technologies (100).

Because of the emerging ethical issues of big data and AI-enabled technologies, a variety of actors (professional organizations, foundations, members of the scientific community, and research institutions) recommended to create new codes of conduct and best-practice guidelines (101). At the international level, too, attempts emerged to update or introduce new data governance rules. The 2018 EU GDPR is one example of this. Another example, in this case related to AI governance, is the 2019 Organization for Economic Co-operation and Development (OECD) “Recommendations on Artificial Intelligence” which offers ethical principles to support AI development (102). Among these principles are inclusiveness, justice, human-centered approach, transparency, quality, and responsibility. In February 2020, the European Commission published a white paper on the same topic, which could also be applicable to the usage of AI for health and clinical purposes and for improving the regulation governing technologies such as AI (103). Regarding the oversight of digital health applications, the World Health Organization (WHO) published a guidance document in 2019 on how to ethically integrate digital health applications into clinical practice while preserving the crucial role of health care professionals

(104). Despite their good intentions, these attempts to reform and improve the governance provide extremely broad principles that are difficult to translate into practice. Furthermore, because of the speed at which these technologies evolve, such recommendations may rapidly become obsolete.

This thesis will analyze the current state of ethical oversight approaches, by examining their gaps. Within the plurality of mechanisms regulating digital health technologies and big data uses for health, this thesis will focus on two of them. First, it will look at the ethical oversight mechanisms regulating health research. The use of big data and AI in health research has already put pressure on RECs and their procedures (105). This trend will only increase as more such studies are conducted in public health and global health. Consequently, more individuals may suffer negative consequences of big data research, thus, the call for a careful ethics review. Second, this thesis will analyze the ethical guidance on digital health applications – specifically for health apps. The reason for this focus on health apps is that they are regulated by elusive guidelines despite being used for a wide range of health, fitness, and well-being related issues (106). As well, health app governance is less strict than that for medical apps and other wearables considered as medical devices. Nonetheless, health apps can collect a variety of data and their use is potentially ethically problematic (107). For example, TraceTogether, the app for the digital contact tracing of COVID-19 cases in Singapore caused a media uproar when it was reported that personal user data collected by the app were shared with the police to assist with criminal investigations (108).

As big data research and health apps become more ubiquitous, there is a need to assess whether the current oversight mechanisms can address the resulting novel ethical issues. A cohesive governance plan could successfully guide stakeholders in their decision-making process. Instead, inconsistent norms and rules would let stakeholders pursue their interests while, potentially, causing harms. Furthermore, the lack of adequate ethical governance could negatively impact data use and sharing. Under these circumstances, to stop or to approve research, as well as whether to promote digital health development would be at the discretion of individual governance mechanisms. This may lead to a fragmented and inconsistent approach that would make collaboration unfeasible within and across jurisdictions. This in turn could result in researchers and developers applying precautionary measures to stop sharing data or circulating new health apps, rather than incurring potential sanctions. In this sense, failing the data governance challenge could ultimately rob society from critical health opportunities and trustworthy technologies.

1.3 Aims and research questions of this thesis

As highlighted so far, ethics and governance are a pressing issue in the age of big data. The emergence of updated regulations and saft-laws signal the intention to advance this conversation. Yet, a variety of ethical considerations remain to be fully explored. Therefore, the first aim of this thesis – which is based on a three-year research study – is to examine the ethical concerns of big data uses and AI-enabled technologies in the field of health research and digital health applications.

Acknowledging the emerging ethical concerns, however, is only the first step toward a more critical appraisal of the big data phenomenon in the health sector. The second step considers how these ethical and technological novelties affect the mechanisms that oversee health research and govern the quality of health apps. This point is crucial as oversight systems determine whether research takes place and whether digital health applications reach the market. Therefore, as a second aim, this thesis investigates whether existing REC mechanisms and health app approval protocols are effective in dealing with the big data ecosystem.

Finally, the third aim is to reflect on how future policies and data governance could tackle the pressing ethical implications of big data and AI-enabled technology in the health sector. The question posed here is whether it is possible, within this rapidly evolving landscape, to provide a stable and reliable governance framework. Given this question, this thesis attempts to define normative recommendations and procedural improvements to tackle the limits of oversight mechanisms. This thesis aspires to pave the way for more ethically aligned big data research and digital health development.

Therefore, the overarching research question of this thesis is:

What are the ethical implications of big data in health research and health apps, and what are the consequences for the existing ethics review mechanisms and data governance more broadly?

As a result of the above considerations, a variety of stakeholders can benefit from this thesis' findings. These include RECs, in Switzerland and abroad; big data researchers and health app developers; and policy makers and legislators working in data governance and ethically aligned digital health. Indirectly, society can profit from a more ethically aligned use of big data and the deployment of AI-enabled technologies.

1.3.1 Structure of the thesis and methods overview

The thesis is comprised of three sections. Each of these three sections explores an interrogative and contributes to answering the main research question presented above.

SECTION I

Research Question:	What are the key ethical concerns of big data uses and AI-enabled technologies in the field of health research and digital health applications?
Objective I:	Map the ethical concerns and clarify their normative value.

SECTION II

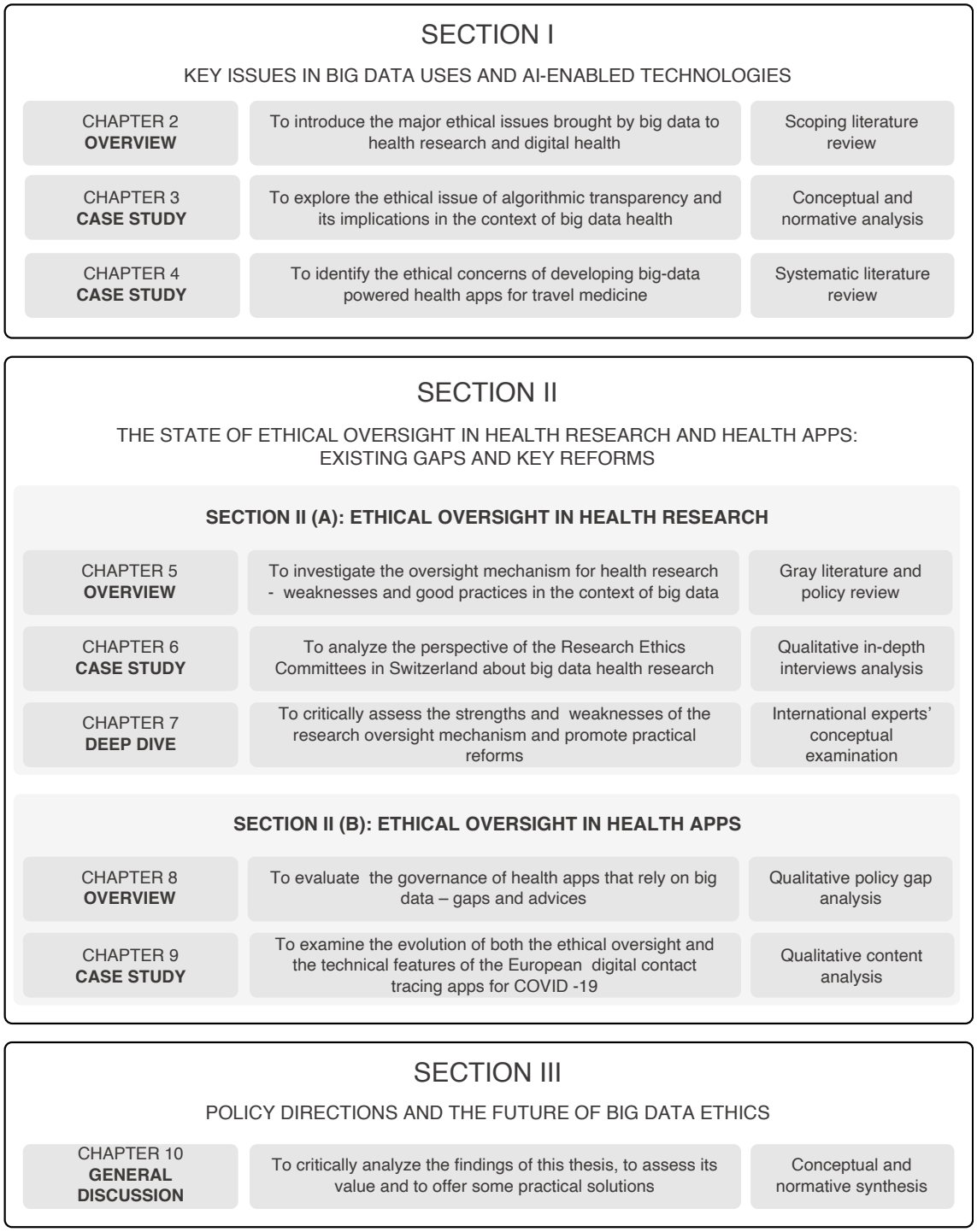
Research Questions:	How do the issues presented in Part I impact the current state of ethical oversight in health research (Section IIA) and health apps (Section IIB)? In other words, which gaps exist in the current oversight mechanisms and what key reforms are needed?
Objective II (A):	Weigh strengths and weaknesses of REC mechanisms judge their level of preparedness in reviewing big data studies, and suggest potential reforms.
Objective II (B):	Review the governance gaps of health apps and suggest strategies to bridge them.

SECTION III

Research Questions:	Which policies and governance should be developed for a future ethically aligned use of big data? What pragmatic solutions could counterbalance the oversight limitations unveiled in Section II?
Objective III:	Offer normative recommendations (to improve governance and policies) and practical suggestions (to orient stakeholders in their decision making).

Including this introduction (Chapter 1), the thesis is composed of ten chapters. Eight of these chapters (Chapters 2-9) are scientific contributions to peer-reviewed journals. Specifically, Chapters 2, 3, 5, 7, 8 have been published, Chapters 4, 6 and 9 are currently under review. Each section of this thesis – excluding Section III – contains chapters classified as either overview, case study or deep dive. The overview chapters summarize the crucial conceptual points of each section. The case study chapters address through specific examples one or more aspects discussed in the overview chapters. The deep dive chapters provide a more profound conceptual analysis, which goes beyond the state-of-the-art description presented in the overview chapters. Figure 1.1 outlines the chapters and their allocation in the various sections, as well as their aims and methodologies.

Figure 1.1: Outline of the chapters



SECTION I: Key ethical issues in big data uses and AI-enabled technologies

In the health sector, the promises of big data boosted by AI systems are increasingly evident. Particularly, the growing literature in this field highlights the potential to improve prevention and treatment, boost clinical research, enhance, and extend access to healthcare, and optimize service delivery. Yet, we should carefully consider the challenges brought about by these unprecedented technologies. Through a scoping review of the literature, *Chapter 2* uncovers a variety of ethical, technical, legal, social, and financial challenges which could affect many stakeholders engaged in health research and digital health development. These include not only researchers and technology developers, but also healthcare workers and the public, which are often the end-users of digital health technologies. This chapter highlights that the research oversight mechanisms (the RECs) may be unprepared for novel big data challenges such as data linkage, biased analytic systems, and data security. RECs could thus consider updating their review processes and criteria, as well as modifying their composition (in terms of members' expertise) to address the challenges and risks that go beyond those of conventional clinical research.

When discussing the big data challenges and risks in health research and health applications, each stage of the data lifecycle should be scrutinized. For instance, some issues might emerge in the data analysis and processing phase, especially if AI systems based on ML methods (i.e., black boxes) are deployed. However, to prevent any ethically problematic outcome coming from these black boxes, we analyze the notion of opacity and assess whether and how opacity can be handled. *Chapter 3* serves as a case study to conceptually explore and distinguish among the different forms of AI opacity, namely lack of disclosure, epistemic opacity, and explanatory opacity, that manifest in digital health technologies. Concrete examples clarify how these forms of opacity emerge, what ethical issues arise from them, and which degree of algorithmic transparency can be achieved through the application of available policies and regulatory tools (e.g., GDPR).

Researchers and developers working on health apps can face a variety of ethical dilemmas; some will be related to AI opacity but not all. In this respect, better health apps arise from assessing - and consequently addressing - the emerging ethical issues. A better-quality app might have a better uptake rate, since the public would perceive it as more trustworthy and acceptable. Hence the use in *Chapter 4* of travel medicine apps as a proxy to identify key ethical considerations that

might affect other health apps too. In this chapter, the relevant literature is systematically searched, and the key ethical issues are examined using thematic content analysis. Privacy, data security and data governance are the most discussed ethical concerns. Despite some papers sporadically mentioning fairness and equity issues, in most cases the literature overlooks issues related to effectiveness, public benefit, user engagement, and risk mitigation, just to mention a few. These findings, however, are not surprising considering the focus of the reviewed papers (i.e., cohort studies and papers describing health app development).

SECTION II (A): The state of ethical oversight in health research: existing gaps and key reforms

Big data and AI are revolutionizing the way researchers conduct biomedical and health research, bringing about novel ethical issues. A variety of stakeholders, such as, research institutions, funding agencies, professional organizations, private companies, and national and international health organizations have issued non-binding guidelines and recommendations in response to these issues. These inconsistent opinions and fragmented guidance confuse those who must assess the ethics of research – namely RECs. *Chapter 5* offers an overview of the challenges experienced by RECs when reviewing big data research and the consequent recommendations to overcome these challenges. Via a scoping-review of the gray literature we collected and analyzed soft-law documents in this domain. The reason for using gray-literature analysis lies in the fact that recommendations, codes of conduct and best practices are often published in the form of soft-law documents, white papers, and technical reports rather than as peer-reviewed research papers. This chapter focuses on whether RECs should have the responsibility of overseeing big data projects given the challenges they face, and, in that case, which strategies can be used to aid them in their task.

None of the documents describing the recommendations for RECs involved the RECs in the norm-development process. This is consistent with the fact that the minority of empirical researchers collected feedback from the RECs about how to best address the big data challenges of health research – and none of it was conducted in Switzerland. Nevertheless, RECs are the ones implementing the ethical guidelines, and therefore should not be considered as only passive recipients. Rather, their involvement may be key in the successful implementation of new rules. To this purpose, *Chapter 6* reports the findings of a qualitative empirical study that conducted semi-structured interviews with the presidents and - when available – the scientific secretaries of all the seven Cantonal RECs in Switzerland. This chapter serves as a case study to learn about

RECs' attitudes, needs, and views on the ethics of big data research. Our results highlight inconsistencies among the RECs about the meaning of big data research, the ethical issues to prioritize, and the solutions to embrace to address these challenges. Conversely, interviewees agreed on appointing ERCs as the key oversight mechanism in charge of big data review, albeit with some caveats. These include training for the committee members on big data and AI challenges and involving external experts when the complexity of the project requires it.

Chapter 7, the final chapter in this section, builds upon the findings of the previous two chapters. This chapter offers a deep dive into the weaknesses of RECs – it explores those limitations that are persistent in RECs, and those which are novel to big data research. The latter ones are further conceptually split between those that characterize the RECs' research purview process, and those related to the RECs' functioning (dependent on the skills, composition, and activities of the RECs). This analysis of weaknesses is the basis for speculating on desirable oversight mechanism improvements. The chapter therefore promotes some reforms. These reforms act on three levels: regulatory, procedural, and complementary. While regulatory reforms involve changes to research ethics regulation, and potentially to international data governance, procedural reforms suggest actionable changes that can be realized by the committees themselves. Complementary reforms, on the other hand, advance the hypothesis for other mechanisms to complement the RECs work. The scope of this chapter is intended to be international, as is the group of authors that contributed to it. In fact, this chapter is anchored in the discussion that emerged during the experts' workshop in "Big Data Challenges for Ethics Review Committees."

SECTION II (B): The state of ethical oversight in health apps: existing gaps and key reforms

While the number and variety of health apps has grown exponentially over the past decade, quality assurance and evaluation has failed to keep pace. Literature suggests that health apps often provide advice that is incomplete, misleading, or wrong, and do not meet the expected standards of privacy and security. Examining the results of the 2017 international census of Data Protection Authorities, *Chapter 8* examines guidance on health apps issued by nine OECD countries as well as the European Commission and the World Health Organization. Through a gap analysis of policy documents, this overview chapter uncovers significant fragmentation in health app governance, as different agencies provide heterogeneous guidance, leaving

developers and consumers to navigate the complex regulatory environment, making compliance onerous and accountability unclear. Therefore, we argue that international coordinated action is urgently needed to guide health app development and use.

The need for such guidance has become even more evident in current circumstances. To help contain the spread of the COVID-19 pandemic, governments worldwide resorted to digital epidemiology tools. In Europe, various stakeholders (i.e., policy makers, technology experts, IT companies and - to a limited extent - the public) collaborated on privacy-preserving health apps for contact tracing. Since their launch, these apps evolved to include new features beyond tracing contacts, including tracking symptoms, checking in at public venues, and counting down quarantine time. *Chapter 9* examines how these apps evolved in eight European countries and whether they follow any ethical guidance. Through qualitative content analysis we show that this evolving trajectory partially echoes the one already seen in Asian countries, although differences persist between the digital health surveillance models. This chapter uses the case of digital contact tracing apps to advance the need for more accurate scrutiny and oversight. As the purpose of digital contact tracing apps expands, privacy should be preserved, and emerging risks monitored.

SECTION III: Policy directions and the future of big data ethics

Finally, *Chapter 10* summarizes and discusses the main findings of the previous sections. As a conclusion to this thesis, the chapter further addresses the thesis' added value by considering its implications for policy, presenting some normative recommendations on how to orient future data governance considering the pressing ethical concerns. Moreover, this conclusion includes some practical suggestions alongside the normative recommendations. These practical suggestions – in the form of an ethical toolkit – could benefit *in primis* RECs, health data researchers, and digital health developers in academia and the private sector. In fact, the toolkit could serve multiple purposes. First, it could raise awareness about the novelties and challenges of big data research. Second, RECs could use the toolkit to preliminarily assess whether they are prepared to review big data research and to fill potential gaps in expertise. Third, the toolkit may guide stakeholders to evaluate whether their health research and products are ethically aligned or not. Such evaluation could be the first step toward preventing harm while promoting valuable research and technologies.

1.3.2 Intellectual property disclosure & ethics statement

This cumulative thesis is the product of collaborative research. AF is the single author of Chapters 1 and 10. However, the scientific contributions presented in Sections I and II have shared intellectual property, and AF is the first author in each of them. Full authorship disclosure is provided at the beginning of each Chapter.

No anticipated physical, psychological, social, or legal risks emerge from this research. The qualitative empirical study presented in Chapter 6 was approved by the ETH Zurich Research Ethics Committee (#EK 2017-N-74).

References

1. Sleep Foundation. Technology in the Bedroom 2020 [updated 2020-11-04. Available from: https://www.sleepfoundation.org/bedroom-environment/technology-in-the-bedroom.
2. International Monetary Fund. World Economic Outlook Database 2015 [Available from: https://www.imf.org/external/pubs/ft/weo/2015/02/weodata/groups.htm#cc.
3. Data Driven Investor. The DotCom Bubble 2021 [updated 2021-01-27. Available from: https://medium.datadriveninvestor.com/the-dotcom-bubble-d33d429addcc.
4. World Economic Forum. Here's how technology has changed the world since 2000 2020 [Available from: https://www.weforum.org/agenda/2020/11/heres-how-technology-has-changed-and-changed-us-over-the-past-20-years/.
5. Clarke T. Creative Destruction, Technology Disruption, and Growth. Oxford Research Encyclopedia of Economics and Finance2019.
6. Garifova L. Infonomics and the Value of Information in the Digital Economy. Procedia economics and finance. 2015;23:738-43.
7. Braidotti R. The posthuman: John Wiley & Sons; 2013.
8. Zuboff S. The age of surveillance capitalism: The fight for a human future at the new frontier of power: Barack Obama's books of 2019: Profile books; 2019.
9. Lupton D. Critical perspectives on digital health technologies. Sociology compass. 2014;8(12):1344-59.
10. Clough PT. The affective turn: Political economy, biomedia and bodies. Theory, Culture & Society. 2008;25(1):1-22.
11. Neff G, Nafus D. Self-tracking: MIT Press; 2016.
12. Eur-Lex. The general data protection regulation 2018 [Available from: https://eur-lex.europa.eu/content/news/general-data-protection-regulation-GDPR-applies-from-25-May-2018.html.
13. World Economic Forum. Why we need to talk about big data 2020 [Available from: https://www.weforum.org/agenda/2020/01/privacy-in-a-world-of-ai-and-big-data/.
14. Goldston D. Big data: Data wrangling. Nature News. 2008;455(7209):15-.
15. The Economist. Data, data everywhere 2010 [Available from: https://www.economist.com/special-report/2010/02/27/data-data-everywhere.
16. Laney D. 3D data management: Controlling data volume, velocity and variety. META group research note. 2001;6(70):1.
17. Ylijoki O, Porras J. Perspectives to definition of big data: a mapping study and discussion. Journal of Innovation Management. 2016;4(1):69-91.
18. De Mauro A, Greco M, Grimaldi M. A formal definition of Big Data based on its essential features. Library Review. 2016.
19. Chollet F. Deep learning with Python: Manning New York; 2018.
20. World Economic Forum. The Fourth Industrial Revolution: what it means and how to respond 2016 [Available from: https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/.
21. Fragidis LL, Chatzoglou PD. Implementation of a nationwide electronic health record (EHR): The international experience in 13 countries. International journal of health care quality assurance. 2018.
22. Ngugi P, Babic A, Kariuki J, Santas X, Naanyu V, Were MC. Development of standard indicators to assess use of electronic health record systems implemented in low-and medium-income countries. Plos one. 2021;16(1):e0244917.

23. Martin-Sanchez F, Verspoor K. Big data in medicine is driving big changes. Yearbook of medical informatics. 2014;9(1):14.
24. Phillips KA, Deverka PA, Hooker GW, Douglas MP. Genetic test availability and spending: where are we now? Where are we going? Health affairs. 2018;37(5):710-6.
25. Genomics England. The 100,000 Genomes Project 2014 [updated 2014-07-21. Available from: https://www.genomicsengland.co.uk/about-genomics-england/the-100000-genomes-project/.
26. Wu P-Y, Cheng C-W, Kaddi CD, Venugopalan J, Hoffman R, Wang MD. –Omic and electronic health record big data analytics for precision medicine. IEEE Transactions on Biomedical Engineering. 2016;64(2):263-73.
27. Vayena E, Haeusermann T, Adjekum A, Blasimme A. Digital health: meeting the ethical and policy challenges. Swiss medical weekly. 2018;148:w14571.
28. MIT Technology Review. More than 26 million people have taken an at-home ancestry test 2019 [Available from: https://www.technologyreview.com/2019/02/11/103446/more-than-26-million-people-have-taken-an-at-home-ancestry-test/.
29. 23andMe. Power of One Million 2015 [updated 2015-06-18. Available from: https://blog.23andme.com/news/one-in-a-million/.
30. Bargh M. Digital health software and sensors: internet of things-based healthcare services, wearable medical devices, and real-time data analytics. American Journal of Medical Research. 2019;6(2):61-6.
31. Elenko E, Underwood L, Zohar D. Defining digital medicine. Nature biotechnology. 2015;33(5):456-61.
32. Shaffer DW, Kigin CM, Kaput JJ, Gazelle GS. What is digital medicine? Studies in health technology and informatics. 2002;80:195-204.
33. Boulos MNK, Brewer AC, Karimkhani C, Buller DB, Dellavalle RP. Mobile medical and health apps: state of the art, concerns, regulatory control and certification. Online journal of public health informatics. 2014;5(3):229.
34. Gambhir SS, Ge TJ, Vermesh O, Spitler R. Toward achieving precision health. Science translational medicine. 2018;10(430).
35. Insel TR. Digital phenotyping: a global tool for psychiatry. World Psychiatry. 2018;17(3):276.
36. Dagum P. Digital biomarkers of cognitive function. NPJ digital medicine. 2018;1(1):1-3.
37. Aramendi AA, Weakley A, Goenaga AA, Schmitter-Edgecombe M, Cook DJ. Automatic assessment of functional health decline in older adults based on smart home data. Journal of biomedical informatics. 2018;81:119-30.
38. Hawgood S, Hook-Barnard IG, O'Brien TC, Yamamoto KR. Precision medicine: beyond the inflection point. Science translational medicine. 2015;7(300):300ps17-ps17.
39. Weber GM, Mandl KD, Kohane IS. Finding the missing link for big biomedical data. Jama. 2014;311(24):2479-80.
40. Jain SH, Powers BW, Hawkins JB, Brownstein JS. The digital phenotype. Nature biotechnology. 2015;33(5):462-3.
41. Vayena E, Gasser U. Strictly biomedical? Sketching the ethics of the big data ecosystem in biomedicine. The ethics of biomedical big data: Springer; 2016. p. 17-39.
42. Noor NLM, Aljunid SA, Noordin N, Teng NIMF, editors. Predictive analytics: the application of J48 algorithm on grocery data to predict obesity. 2018 IEEE Conference on Big Data and Analytics (ICBDA); 2018: IEEE.
43. Ophir Y, Tikochinski R, Asterhan CS, Sisso I, Reichart R. Deep neural networks detect suicide risk from textual facebook posts. Scientific reports. 2020;10(1):1-10.
44. Forbes. How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did 2012 [Available from: https://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/.

45. Sheth A, Jaimini U, Thirunarayan K, Banerjee T, editors. Augmented personalized health: how smart data with IoTs and AI is about to change healthcare. 2017 IEEE 3rd International Forum on Research and Technologies for Society and Industry (RTSI); 2017: IEEE.
46. National Institute of Health. All of Us Research Program Expands Data Collection Efforts with Fitbit 2019 [updated Thu, 08/20. Available from: https://www.ncbi.nlm.nih.gov/pubmed/.
47. Lopreite M, Panzarasa P, Puliga M, Riccaboni M. Early warnings of COVID-19 outbreaks across Europe from social media. Scientific Reports. 2021;11(1):2147.
48. AMA J Ethics. How Will Artificial Intelligence Affect Patient-Clinician Relationships? : 22(5):E395-400; 2020 [Available from: https://journalofethics.ama-assn.org/article/how-will-artificial-intelligence-affect-patient-clinician-relationships/2020-05.
49. The Harvard Gazette. Risks and benefits of an AI revolution in medicine 2020 [updated 2020-11-11. Available from: https://news.harvard.edu/gazette/story/2020/11/risks-and-benefits-of-an-ai-revolution-in-medicine/.
50. Obermeyer Z, Emanuel EJ. Predicting the future—big data, machine learning, and clinical medicine. The New England journal of medicine. 2016;375(13):1216.
51. Beam AL, Kohane IS. Big data and machine learning in health care. Jama. 2018;319(13):1317-8.
52. Wernick MN, Yang Y, Brankov JG, Yourganov G, Strother SC. Machine learning in medical imaging. IEEE signal processing magazine. 2010;27(4):25-38.
53. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. nature. 2017;542(7639):115-8.
54. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. Jama. 2016;316(22):2402-10.
55. Chen J, Wu L, Zhang J, Zhang L, Gong D, Zhao Y, et al. Deep learning-based model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography. Scientific reports. 2020;10(1):1-11.
56. Rezaii N, Walker E, Wolff P. A machine learning approach to predicting psychosis using semantic density and latent content analysis. NPJ schizophrenia. 2019;5(1):1-12.
57. Ming C, Viassolo V, Probst-Hensch N, Dinov ID, Chappuis PO, Katapodi MC. Machine learning-based lifetime breast cancer risk reclassification compared with the BOADICEA model: impact on screening recommendations. British journal of cancer. 2020:1-8.
58. Filipp FV. Opportunities for artificial intelligence in advancing precision medicine. Current genetic medicine reports. 2019;7(4):208-13.
59. Murray CJ, Alamro NMS, Hwang H, Lee U. Digital public health and COVID-19. The Lancet Public Health. 2020;5(9):e469-e70.
60. Sheath D, Ruiz De Castaneda RL, Bempong N-E, Raviglione M, Machalaba C, Pepper MS, et al. Precision global health: a roadmap for augmented action. Journal of Public Health and Emergency. 2020;4(5):1-12.
61. European Commission. EPIWORK: Ebola forecasting uses model developed by EU project 2014 [Available from: https://ec.europa.eu/programmes/horizon2020/en/news/epiwork-ebola-forecasting-uses-model-developed-eu-project.
62. Hussain A, Tahir A, Hussain Z, Sheikh Z, Gogate M, Dashtipour K, et al. Artificial intelligence-enabled analysis of UK and US public attitudes on Facebook and Twitter towards COVID-19 vaccinations. medRxiv. 2020.

63. Merz AA, Gutiérrez-Sacristán A, Bartz D, Williams NE, Ojo A, Schaefer KM, et al. Population attitudes toward contraceptive methods over time on a social media platform. American Journal of Obstetrics and Gynecology. 2020.
64. Wilton Park. Reimagining global health: self-care interventions and implications for healthcare. 2018 [Available from: https://www.wiltonpark.org.uk/wp-content/uploads/WP1639-Report.pdf.
65. Vayena E, Ferretti A. Big Data and Artificial Intelligence for Global Health. Global Health: Ethical Challenges. 2020:429.
66. Wahl B, Cossy-Gantner A, Germann S, Schwalbe NR. Artificial intelligence (AI) and global health: how can AI contribute to health in resource-poor settings? BMJ global health. 2018;3(4):e000798.
67. Vayena E, Dzenowagis J, Brownstein JS, Sheikh A. Policy implications of big data in the health sector. Bulletin of the World Health Organization. 2018;96(1):66.
68. Sharon T. When digital health meets digital capitalism, how many common goods are at stake? Big Data & Society. 2018;5(2):2053951718819032.
69. Apple. Apple announces three groundbreaking health studies 2019 [Available from: https://www.apple.com/newsroom/2019/09/apple-announces-three-groundbreaking-health-studies/.
70. Morley J, Taddeo M, Floridi L. Google health and the NHS: overcoming the trust deficit. The Lancet Digital Health. 2019;1(8):e389.
71. Leonelli S. Scientific Research and Big Data: Edward N. Zalta (ed.); 2020 [updated 2020-05-29. Available from: https://plato.stanford.edu/entries/science-big-data/#BigDataRiskEthiDataScie.
72. O'Brien D, Ullman J, Altman M, Gasser U, Bar-Sinai M, Nissim K, et al. Integrating Approaches to Privacy Across the Research Lifecycle: When Is Information Purely Public? Berkman Center Research Publication. 2015(2015-7).
73. Norval C, Henderson T. Contextual consent: Ethical mining of social media for health research. arXiv preprint arXiv:170107765. 2017.
74. Vayena E, Madoff L. Navigating the Ethics of Big Data in Public Health. The Oxford Handbook of Public Health Ethics. 2019:354.
75. Nissenbaum H. Privacy as contextual integrity. Wash L Rev. 2004;79:119.
76. Holm S, Ploug T. Big data and Health Research—the governance challenges in a mixed data economy. Journal of bioethical inquiry. 2017;14(4):515-25.
77. Mittelstadt BD, Floridi L. The ethics of big data: current and foreseeable issues in biomedical contexts. The ethics of biomedical big data. 2016:445-80.
78. Hayden EC. Informed consent: a broken contract. Nature. 2012;486(7403).
79. Shilo S, Rossman H, Segal E. Axes of a revolution: challenges and promises of big data in healthcare. Nature medicine. 2020;26(1):29-38.
80. Hariri RH, Fredericks EM, Bowers KM. Uncertainty in big data analytics: survey, opportunities, and challenges. Journal of Big Data. 2019;6(1):1-16.
81. Vayena E, Blasimme A. Biomedical big data: new models of control over access, use and governance. Journal of bioethical inquiry. 2017;14(4):501-13.
82. Char DS, Shah NH, Magnus D. Implementing machine learning in health care—addressing ethical challenges. The New England journal of medicine. 2018;378(11):981.
83. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence. 2019;1(5):206-15.
84. Blasimme A. Producing health: Georges Canguilhem and the contemporary clinical discourse about aging. The Normal and the Pathological: Essays on Georges Canguilhem: Hermann; 2020.

85. Vayena E, Gasser U, Wood AB, O'Brien D, Altman M. Elements of a new ethical framework for big data research. Washington and Lee Law Review Online. 2016.
86. Xafis V, Schaefer GO, Labude MK, Brassington I, Ballantyne A, Lim HY, et al. An ethics framework for big data in health and research. Asian Bioethics Review. 2019;11(3):227-54.
87. Ajunwa I, Crawford K, Ford JS. Health and big data: An ethical framework for health information collection by corporate wellness programs. The Journal of Law, Medicine & Ethics. 2016;44(3):474-80.
88. Miller B. Is Technology Value-Neutral? Science, Technology, & Human Values. 2021;46(1):53-80.
89. Franssen M, Lokhorst G-J, van de Poel I. Philosophy of Technology 2009 [updated 2009-02-20. Available from: https://plato.stanford.edu/entries/technology/#AnalPhilTech.
90. Wheeler M. Martin Heidegger 2011 [updated 2011-10-12. Available from: https://plato.stanford.edu/entries/heidegger/.
91. Winner L. Autonomous technology: Technics-out-of-control as a theme in political thought: Mit Press; 1978.
92. Tempini N. Till data do us part: Understanding data-based value creation in data-intensive infrastructures. Information and Organization. 2017;27(4):191-210.
93. Vayena E. Value from health data: European opportunity to catalyse progress in digital health. The Lancet. 2021;397(10275):652-3.
94. Price W, Nicholson I. Black-box medicine. Harv JL & Tech. 2014;28:419.
95. Lupton D. Self-tracking modes: Reflexive self-monitoring and data practices. Available at SSRN 2483549. 2014.
96. Morley J, Floridi L. The limits of empowerment: how to reframe the role of mHealth tools in the healthcare ecosystem. Science and engineering ethics. 2019:1-25.
97. McMahon A, Buyx A, Prainsack B. Big data governance needs more collective responsibility: the role of harm mitigation in the governance of data use in medicine and beyond. Medical law review. 2020;28(1):155-82.
98. Forbes. Are Research Ethics Obsolete In The Era Of Big Data? 2016 [Available from: https://www.forbes.com/sites/kalevleetaru/2016/06/17/are-research-ethics-obsolete-in-the-era-of-big-data/.
99. Vayena E, Salathé M, Madoff LC, Brownstein JS. Ethical challenges of big data in public health. PLoS Comput Biol. 2015;11(2):e1003904.
100. Sharon T, Lucivero F. Introduction to the Special Theme: The expansion of the health data ecosystem–Rethinking data ethics and governance. SAGE Publications Sage UK: London, England; 2019.
101. Jobin A, Ienca M, Vayena E. The global landscape of AI ethics guidelines. Nature Machine Intelligence. 2019;1(9):389-99.
102. Organization for Economic Co-operation and Development. Recommendation of the Council on Artificial Intelligence 2019 [Available from: https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449.
103. European Commission. White Paper on Artificial Intelligence: a European approach to excellence and trust 2020 [Available from: https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en.
104. Labrique A, Agarwal S, Tamrat T, Mehl G. WHO Digital Health Guidelines: a milestone for global health. npj Digital Medicine. 2020;3(1):1-3.
105. Sellers C, Samuel G, Derrick G. Reasoning “uncharted territory”: notions of expertise within ethics review panels assessing research use of social media. Journal of Empirical Research on Human Research Ethics. 2020;15(1-2):28-39.

106. Bates DW, Landman A, Levine DM. Health apps and health policy: what is needed? Jama. 2018;320(19):1975-6.
107. Huckvale K, Prieto JT, Tilney M, Benghozi P-J, Car J. Unaddressed privacy risks in accredited health and wellness apps: a cross-sectional systematic assessment. BMC medicine. 2015;13(1):1-13.
108. BBC News. Singapore reveals Covid privacy data available to police 2021 [Available from: https://www.bbc.com/news/world-asia-55541001.

SECTION I

KEY ISSUES IN BIG DATA USES
AND AI-ENABLED TECHNOLOGIES

CHAPTER 2: Considerations for Ethics Review of Big data Health Research: a Scoping Review

Peer reviewed and published as: Ienca, M.*, Ferretti, A.*, Hurst, S., Puhan, M., Lovis, C., Vayena, E. (2018). Considerations for ethics review of big data health research: A scoping review. *PLOS ONE*, 13(10), e0204937, doi:10.1371/journal.pone.0204937.

* These authors contributed equally and share first authorship

2.1 Abstract

Big data trends in biomedical and health research enable large-scale and multi-dimensional aggregation and analysis of heterogeneous data sources, which could ultimately result in preventive, diagnostic and therapeutic benefit. The methodological novelty and computational complexity of big data health research raises novel challenges for ethics review. In this study, we conducted a scoping review of the literature using five databases to identify and map the major challenges of health-related big data for Ethics Review Committees (ERCs) or analogous institutional review boards. A total of 1093 publications were initially identified, 263 of which were included in the final synthesis after abstract and full-text screening performed independently by two researchers. Both a descriptive numerical summary and a thematic analysis were performed on the full texts of all articles included in the synthesis. Our findings suggest that while big data trends in biomedicine hold the potential for advancing clinical research, improving prevention and optimizing healthcare delivery, yet several epistemic, scientific and normative challenges need careful consideration. These challenges have relevance for both the composition of ERCs and the evaluation criteria that should be employed by ERC members when assessing the methodological and ethical viability of health-related big data studies. Based on this analysis, we provide some preliminary recommendations on how ERCs could adaptively respond to those challenges. This exploration is designed to synthesize useful information for researchers, ERCs and relevant institutional bodies involved in the conduction and/or assessment of health-related big data research.

2.2 Introduction

The generation of digital data has drastically increased in the last years due to the ubiquitous deployment of digital technology as well as advanced computational analytics techniques (1, 2). The term big data is still vaguely defined. In general terms, big data involves large sets of data with diverse levels of analysable structuration, coming from heterogeneous sources (online data, social media profiles, financial records, self-tracked parameters, etc.), produced with high frequency and which can be further processed and analysed using computational techniques. While the term big data has become nearly ubiquitous, there is controversy over what data volumes are sufficiently large to obtain the big data label. Dumbill, for example, suggested that

data should be considered big when they cross the threshold of the conventional databases systems' capacity in processing information (3).

Big data trends characterize various sectors including basic science (1, 4), business (5), government (6), national security (7) and transportation (8). Big data trends have increasingly pervaded also the healthcare domain, as new health-related data sources have grown in volume and variety, and became available for large-scale aggregation and high-speed analysis (9). These include Electronic Health Records (EHRs), data from mobile health (mHealth) applications, medical blogs and web-networks (10) (11), healthcare robotics (12), medical internet of things (13), as well as direct-to-consumer genetic (14), and screening tests (15). Additionally, health-related information can be derived not only from digital health applications, but also from non-strictly-medical data sources (16) such as online personal dietary programs, fitness club memberships and Twitter hashtags (17). Health-related big data is the umbrella term used to describe extremely large and heterogeneous data sets that may be analysed computationally to reveal patterns, trends, and correlations, that have relevance for human health (18).

The availability of health-related big data holds the promise of exerting a positive impact on biomedical research. For example, tailoring diagnostics to automated analyses of high resolution images has become a standard procedure in cancer research (19). In parallel, mapping and collecting large-scale data volumes enables the creation of epidemiological models that can inform about an epidemics' space-time propagation. Finally, novel and patient-tailored therapeutic opportunities might emerge from the possibility of continuously monitoring patient health, tracking pathologic characteristics at specific points in time, and aggregating heterogeneous data sources (20). These benefits might occur both in public health and at the individual level. Bates (21) argued that the use of big data has a valuable impact on public health, since it might help identify and promptly intervene on high-risk and high-cost patients.

While opening the prospect of clinical benefit, the use of health-related big data raises important challenges. In light of their methodological novelty, potentially far-reaching impacts, and computational complexity, big data approaches to human health have been claimed to raise ethical, legal and social implications (22). Ethical and legal challenges include the risk to compromise privacy, personal autonomy, and the solidarity-based approach to healthcare funding, as well as effects on public demand for transparency, trust, and fairness while using big data (23). Furthermore, authors have listed data heterogeneity, data protection, analytical flows

in analysing data, and the lack of appropriate infrastructures for data storage as critical technical and infrastructural issues that might endanger a big-data-driven healthcare (24). While some of these challenges have received scientific and institutional attention, other ones have remained largely unexplored. In 2016, a review identified a number of areas of concern associated with health-related big data that did not obtain adequate attention among researchers (22). These included group-level ethical harms, the intimate link between epistemological and ethical issues, the distinction between harms to data subject resulting from, respectively, academic and commercial uses of big data, the problematic fiduciary relationship between data custodian and data subjects, the role of data ownership and intellectual property as a mechanism for data control, and, finally, the provision of data access rights to data subjects.

The ethical, legal and social implications of health-related big data raise novel challenges also for Ethics Review Committees (ERCs). ERCs and institutional review boards are increasingly requested to evaluate an ever-growing number of research projects and associated activities involving big data (large data volumes and big data analytics), whose risks and benefits often appear hard to assess. Some authors have called for the development of comprehensive regulatory policies for healthcare entities and new computing safeguards that can address public concerns, such as the protection of individually identifiable information (25). However, in absence of specific guidelines and comprehensive evaluation studies, ERCs might be facing uncertainty on how to review health-related big data projects and according to which evaluative criteria. In fact, researchers have observed that traditional conceptual tools and/or legal requirements for ethics review in clinical research like informed consent, minimal risk and fair subject selection might be of limited help, if not ill suited, for the evaluation of big data projects (26, 27). The reason for that stems from the fact that these tools were conceived in the context of conventional clinical research (e.g. clinical trials) not in connection to the evolving applications and innovative research designs of big data research (27). For example, informed consent is often not practical to obtain for studies involving a retrospective access to data from millions of individuals.

The nature of big data studies also challenges the current mandate and purview of ERCs. For example, studies involving publicly available and anonymized data have traditionally been perceived to be outside of the purview of ERCs. This would include data from Twitter (which are public by default), Facebook or other online platforms. Furthermore, ethical safeguards for human subjects research “are often written with definitions that exclude Internet research” (28). This is problematic for a twofold reason. First, research has shown that big data analytics can reveal

sensitive information from seemingly innocuous public data points, including information that the original data generators might reasonably wish to keep private. For example, a recent study has successfully used deep neural networks to predict the sexual orientation of users based on facial images from public profiles posted on dating website (29). Second, several studies have shown that de-identified (30) and even anonymized data (31) can be reverse engineered to re-identify individuals, leading experts to the conclusion that “there is no such thing as anonymous data”. This raises the question of whether big data projects should require oversight by an ERC even when the data collected are public and anonymized or de- identified. A recent systematic review has concluded that most normative documents deem the review of an ERC as necessary to address the concerns associated with the use of anonymized data for research (32). In contrast, when ERCs waived the review of big data studies involving publicly available and anonymized data repositories because they considered them outside their purview, such as in the case of Facebook’s “emotional contagion” study (33), experts criticized this narrow interpretation of the ERC’s mandate (34).

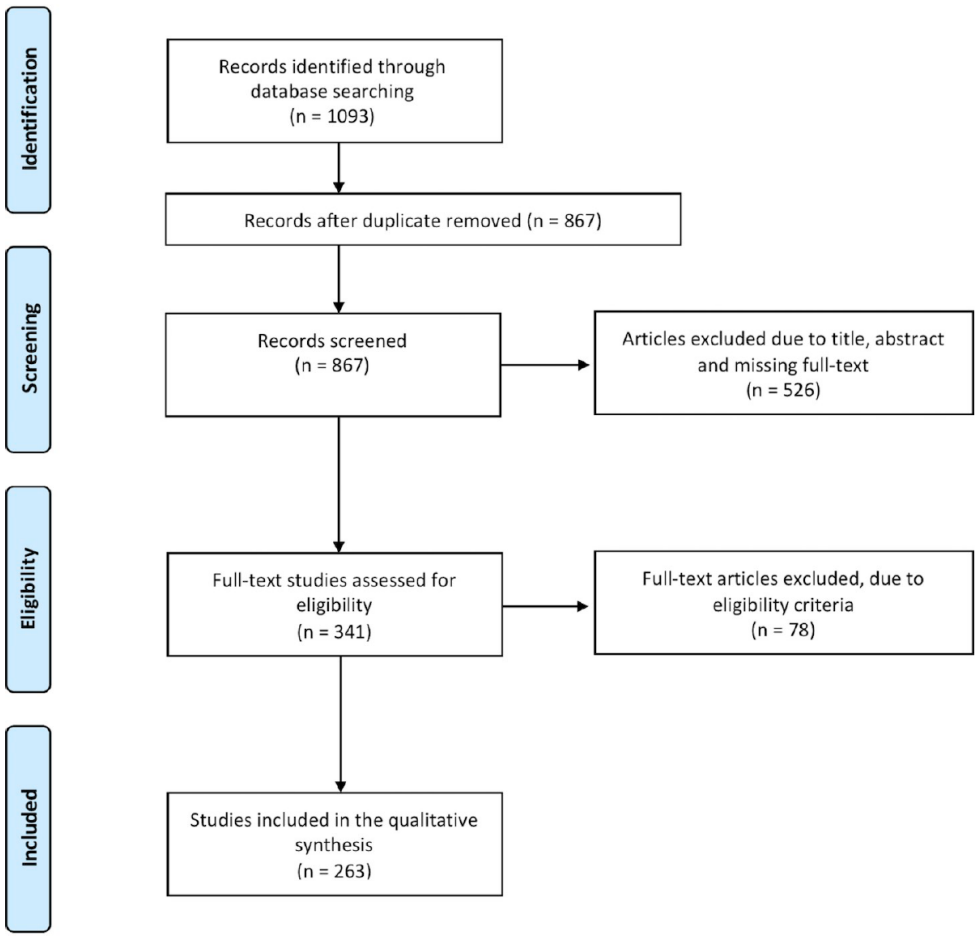
In the present study, we aim to identify the promises and challenges of health-related big data research that have relevance for ERCs. Furthermore, we use these findings to suggest how ERCs could adaptively respond to this methodological transformation. This exploration is designed to synthesize useful information for researchers, ERCs and relevant institutional bodies involved in the conduction and/or assessment of health-related big data research.

2.3 Methods

On the 18th of September 2018 we conducted a scoping review of the scientific literature and searched five databases (EMBASE, Web of Science, Pubmed, IEEE Xplore, and Scopus) to retrieve eligible publications. We searched title, abstract, and keywords for the terms: ("big data" OR “Artificial Intelligence” OR "data science" OR "digital data") AND (“medical” OR “healthcare” OR “clinical” OR "personalised medicine") AND (“policy” OR “ethics” OR “governance” OR "ethics committee" OR “IRB” OR "review board" OR “assessment”). Query logic was modified to adapt to the language used by each engine or database (Appendix 1 Chapter 2). Screening identified 1093 entries. All entries were imported into the Endnote literature manager software. Three phases of filtering were performed independently by two researchers to minimize subjective bias.

The scoping review is a review method aimed at synthesizing research evidence and mapping the existing literature in a certain field of interest (35). Unlike a systematic review, scoping review methods are considered of particular use when the topic has not yet been extensively reviewed or is of a complex or heterogeneous nature (35, 36). Following the recommendations by Pham et al. (36), the study selection process was conducted and presented using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (http://prisma-statement.org/) as a guide (see Figure 2.1).

Figure 2.1: Scoping literature review flow chart (PRISMA)



First, duplicates were removed both automatically using the Endnote tool for duplicate detection and manually based on abstract screening. A total of 226 articles was removed at this stage.

Second, eligibility assessment was performed independently by two of the co-authors on the remaining 867 articles through title- abstract screening and, subsequently, full text screening.

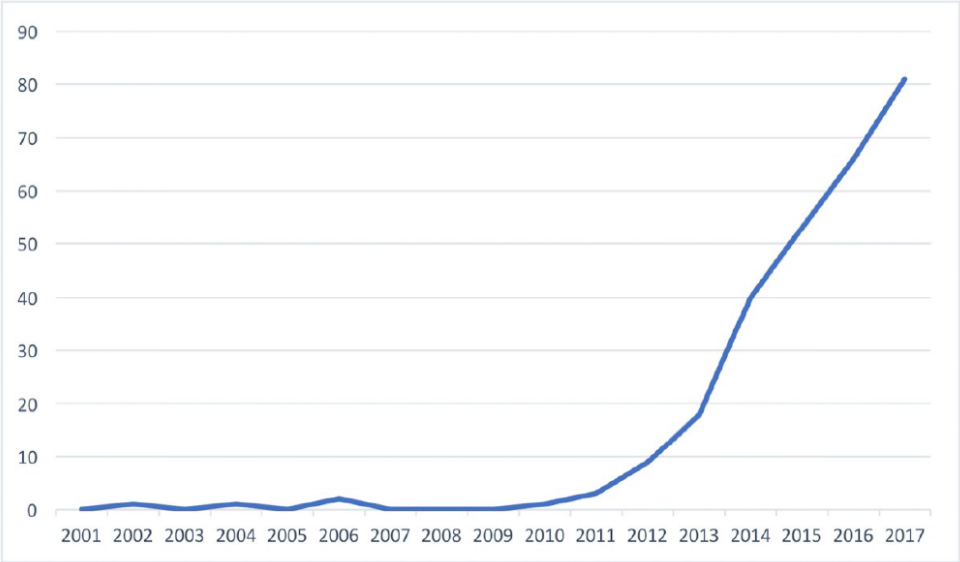
Diverging inclusion choices between the two reviewers were discussed with the research group with documented reasons. Studies included in the synthesis had the following features: (i) original articles, book chapters or conference proceedings; (ii) written in English, Italian, French or German (languages spoken by the researchers); (iii) published before September 18th, 2017; and (iv) focused on the assessment of big data trends in the biomedical/healthcare context. Reviews, letters to the editors, business reports and dissertations were not included. A total of 263 studies were included in the final synthesis and imported manually into Microsoft Excel 15.40 format based on a shared data-charting form. Following the recommendations to enhance scoping study methodology delineated by Levac et al. (37), the data-charting form was collectively developed by our research team to determine which variables to extract from the review data.

Third, based on the same recommendations, we performed both a descriptive numerical summary and a thematic analysis. In the former analysis, both relative and cumulative frequencies were extracted and graphically represented using bar charts. Following Arksey and O'Malley (36), our descriptive numerical summary also included the total number of articles included, types of study design (empirical vs. non empirical), years of publication etc. In the latter analysis, recurrent thematic patterns were identified through full-text screening and subsequent coding. The coding phases was independently performed by two researchers. Once conceptually stable thematic patterns emerged from the codes, these were grouped together into a system of themes and subthemes. All entries were checked anew through an automated text search for the presence of the emerging themes. Following Braun and Clarke (38), codes that did not seem to fit into any main theme, were temporarily housed in a “miscellaneous” group and subsequently either clustered into a new theme or reallocated to an existing thematic group after consultation. Internal consultation was performed among all members of our research team to integrate and validate our findings.

2.4 Results

Our results reveal a large, diverse and rapidly growing body of literature on the impact of big data in the biomedical domain. Data show that the overall number of articles published in the time period 2012–2017 is 131 times higher compared to the period 2001–2005 as represented in Figure 2.2.

Figure 2.2: Increase over time in research papers discussing the challenges of health-related big data.
N.B. The search was performed on September 18, 2017. Therefore, the full number of articles for year 2017 was calculated by projecting the data until September 18.



Data breakdown by medical speciality and field of medical application indicates that big data approaches have been discussed and evaluated in relation to several branches of medicine including neurology and psychiatry (n = 31), oncology (n = 17), cardiology (n = 8), medical genetics (n = 8), immunology and infectious diseases (n = 8), as well as nuclear medicine and radiology (n = 6). Crossfield evaluations of health-related big data appeared highly prevalent (n = 155).

Thematic analysis identified a number of potential opportunities and challenges associated with health-related big data approaches, many of which have relevance for ethics review. Opportunities could be grouped into four main themes: biomedical research, prevention, healthcare delivery and healthcare management. Potential benefits in the research domain include the possibility of collecting real-world data, accelerating the development of new medical technology and facilitating translational research. Big data was also associated with the improvement of preventive measures at both the individual and population level. In relation to care delivery, the following benefits were envisioned: precision and personalized medicine, earlier and more accurate diagnostics, enhanced clinical decision-making, ubiquitous health monitoring,

improved patient safety and better therapy. Subsequent numeric analysis of thematic clusters is presented in Table 2.1.

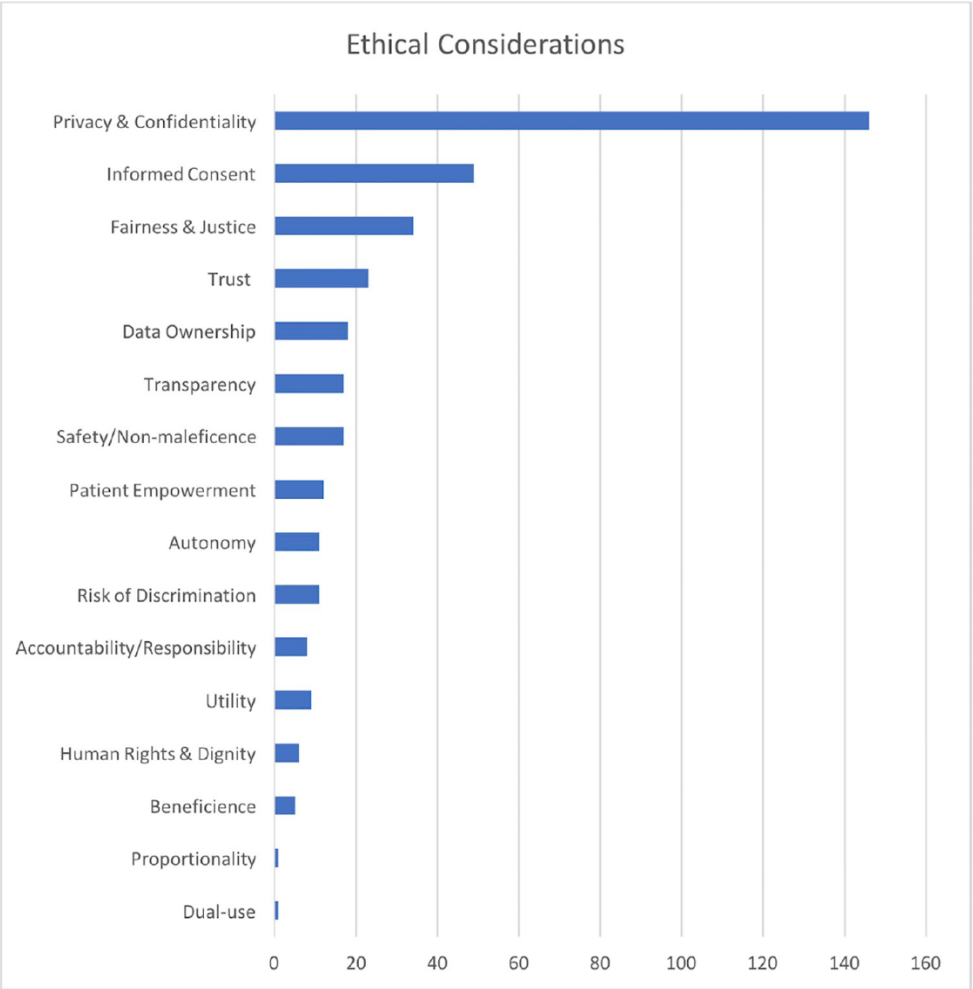
Table 2.1: Recurrent promises and challenges associated with health-related big data that have relevance for ethics review.
N.B. The same study might describe >1 promise or challenge.

Opportunities	Challenges
Healthcare Delivery (n = 276)	Technical (n = 125)
Healthcare Management (n = 90)	Ethical (n = 81)
Biomedical Research (n = 85)	Methodological (n = 66)
Prevention (n = 45)	Regulatory (n = 39)
	Social (n = 16)
	Infrastructural (n = 11)
	Financial (n = 10)

Envisioned challenges appeared of seven major types: technical (n = 125), ethical (n = 81), methodological (n = 66), regulatory (n = 39), social (n = 16), infrastructural (n = 11) and financial (n = 10). Technical challenges relate to issues inherent in the data ecosystem. These include data security, data quality, data storage, data linkage, and tools for data reuse. Methodological challenges relate to the system of methods used in the study and include issues of standardizing data and metadata, integrating and processing data, monitoring resource utilisation and compensating for incomplete data. Regulatory challenges relate to rules or directives such as those regulating data ownership and the accountability of actors in relation to the potential risks associated with using and managing data. Social challenges are those that have relevance for human society and its members. These include, among others, secondary uses of data in relation to participants consent, sociocultural and ethnic bias and subsequent risk of discrimination, power asymmetries between data subjects and data controllers. Finally, financial and infrastructural issues included the financial viability of data storage sites and to the level of preparedness of existing infrastructures respectively.

Ethical challenges are those related to moral principles. Our analysis revealed privacy and confidentiality to be by far the dominant concern (n = 146) in the ethical domain, followed by informed consent (n = 49), fairness and justice (n = 34), trust (n = 23), data ownership (n = 18) and others. Figure 2.3 presents a full overview of ethical considerations associated with health-related big data studies with associated relative frequencies.

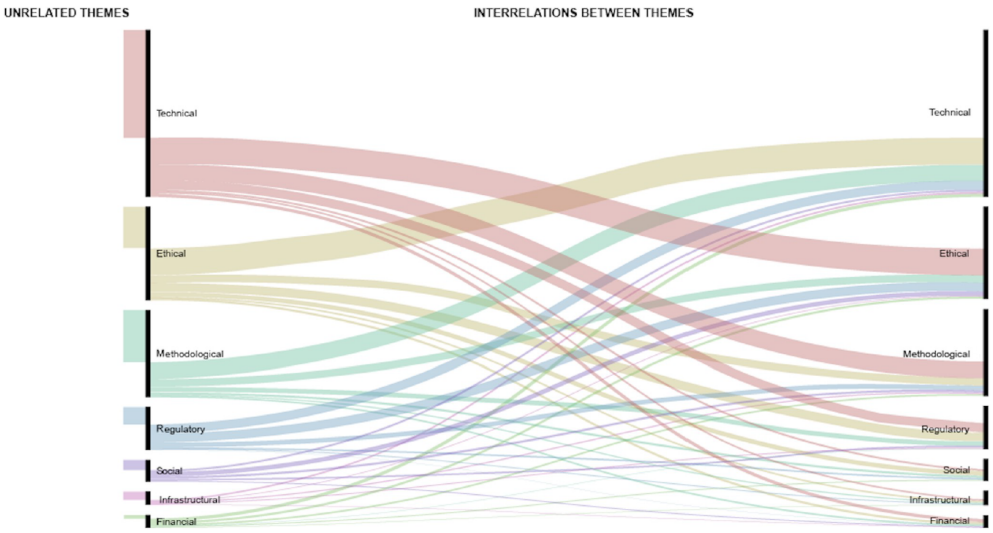
Figure 2.3: Frequency of ethical considerations associated with health-related big data studies



While the analysis revealed a number of implications with relevance for ethics review, only 13% of reviewed studies provided specific normative recommendations for ERCs or other institutional review boards. Data breakdown by study methodology revealed that only a small portion of those recommendations (n = 5; 14%) was informed by empirical methods.

A subsequent analysis of thematic co-occurrences shows a strong mutual relationship between different thematic families, especially between technical and ethical issues, as shown in Figure 2.4. In particular, technical issues such as data security and data linkage were often presented in coordination with ethical issues such as personal privacy.

Figure 2.4: Alluvial diagram of mutual interrelations between different thematic families (figure credit Joanna Sleigh)



2.5 Limitations

This study presents four main limitations. First, a selection bias might be present since the search retrieved only articles written in languages known by the researchers (English, French, German and Italian), excluding articles written in other languages. A similar limitation affects database selection as searching other databases may have possibly identified additional relevant studies. While this risk of selection bias applies to any review since the number of databases that can be feasibly searched is always finite, we attempted to minimize selection bias by exploring both domain-general and domain-specific databases, including the major databases in biomedical research and computer science, which represent the primary interdisciplinary intersection when it comes to biomedical big data. Second, as it was often observed in relation to scoping reviews, the explorative nature and broad focus of our search methodology makes it 'unrealistic to retrieve and screen all the relevant literature' (39). However, one advantage of the scoping methodology is the opportunity to explore also the grey literature and the secondary sources (e.g. bibliographies of retrieved papers), which is likely to increase comprehensiveness. The breadth of the research focus might have inevitably affected the depth of the analysis. The reason for that stems from the fact that the outcomes of a scoping review, compared to systematic review methods, are "more narrative in nature" (40) and usually not presented through descriptive

statistical analysis. Finally, our review included very heterogeneous studies and did not assess the study quality. The reason for that stems from the fact that our main goal was to explore the entire range of challenges that have relevance for ERCs, regardless of how those challenges were originally addressed and discussed. While these four limitations might prevent generalization, we believe that the scoping methodology was best suited to reflect the explorative nature and broad focus of our research question. In fact, it has often been noted, that scoping reviews are not intended to be exhaustive (41, 42) or to provide detailed statistical analyses (40) but to map an heterogeneous body of literature related to a broad and novel topic (35). As scoping reviews are usually considered a "richly informed starting point for further investigations" (40), future studies should consider this work as a preliminary step to a systematic review and associated statistical data analysis. Furthermore, they could use this general mapping of the health-related big data topic to generate empirically testable research hypotheses.

2.6 Discussion

The drastic increase over the past 5 years in the number of studies discussing the implications of health-related big data confirms the research community's increasing attention to the applicability of big data approaches into the healthcare domain. As the application of big data in healthcare (43) and the market size forecasts for big data hardware, software and professional services investments in the healthcare and pharmaceutical industry are growing steadily (44), there will be a parallel need to assess the impact of this expanding sociotechnical trend. This expansion can be seen as a sign of what has been defined the "inevitable application of big data to healthcare" (10) induced by the widespread uptake of electronic health records (EHRs), and the large-scale storing and sharing of genomic, proteomics, imaging and many other biomedical data.

The large prevalence of cross-field evaluations of health-related big data is an indicator of the potential of big data approaches to aggregate data from multiple medical data sources (e.g. combining data about gene expression and brain function in neurogenic studies) and multiple levels of clinical intervention (e.g. linking prevention and diagnostics to therapy and care delivery). In addition, analyses show that clinical outcomes can be produced from novel and non-strictly medical data sources. These include using Twitter to track and even forecast disease activity (45), exploiting Facebook data for suicide prevention (46), or using seasonal pollen forecast to predict asthma (47, 48). On the long term, this meta-specialty nature of big data approaches is

likely to blur traditional separations between different medical specialties and levels of clinical intervention, opening more interfaces for inter- specialty exchange in the healthcare and biomedical research domains. This will raise the challenge for ERCs to review big data projects without relying on traditional discrete taxonomies of medical specialization and/or models of clinical application. In parallel, our findings illustrate the potential applicability of big data approaches to an increased variety of medical specialties. While branches of medicine like oncology (49, 50), radiology (51) and clinical genetics (52) were already known to be particularly suitable for big data approaches, our review revealed a promising outlook associated with using big data in several other medical domains including neurology (53, 54), psychiatry (55), immunology (56), nephrology (57), and geriatrics (58).

The high frequency of technical challenges addressed when assessing health-related big data highlights the persistence of a number of technical weaknesses and limitations, most of which are likely dependent on the historical novelty of such sociotechnical trend. These include problems of data quality, integrity, and security. Developing robust technical solutions that can guarantee the quality, integrity and security of the data, and allow their secure transmission, linkage and storage, was often presented as a priority for any successful deployment of big data for human health. This might require the development of better security-protecting infrastructures, data wrangling and scripting (e.g. batch processing) tools for data cleansing in order to guarantee the quality of data -for example, through automatic detection and removal of corrupt or inaccurate records- as well as techniques that can preserve the integrity of data through the entire data cycle, prevent corruption and enable interoperability. Furthermore, distributed ledger technology, distributed storage and incremental analytics are also believed to hold promises in the health domain (59, 60). From the perspective of ERCs, this implies a more rigorous yet systemic oversight (61) of technical considerations to guarantee that the afore listed safeguards are implemented by the researchers.

The relative frequency of methodological issues, however, highlights that fixing technical problems alone might not be sufficient to use big data for good. ERCs are usually required to evaluate the methodological soundness of a study if this has ethical consequences. For example, if a RCT is designed without giving participants an equal chance of being assigned to any group, ERCs are entitled to assess the methodological soundness of the study to preserve the principle of fairness. For the same reason, in the context of big data research, ERCs might be entitled to assess the soundness of studies whose methods may result in algorithmic discrimination or breaches of personal privacy. For example, they may examine whether the researchers have

implemented all necessary safeguards to prevent algorithmic bias and comply with data security standards.

Examining the methodological soundness of health-related big data studies will likely require the adoption of different assessment criteria compared to traditional biomedical research. For example, it may require a rethinking of what counts as “public” data and what counts as “harm” in data-driven research. In addition, big data research is usually not based on the formulation and testing of specific research hypotheses, but on the identification of patterns from large volumes of data. This hypothesis-free nature of (some) big data research makes it harder to apply conventional epistemological mechanisms for scientific demarcation and quality control like falsifiability and refutability (62). This poses for ERCs the problem of clearly demarcating the explanatory power of big data driven research. Researchers have questioned that big data analytics might speak for themselves (63) independent of explanatory hypotheses and refuted the idea that they can be used for biomedical purposes in absence of robust and causally explanatory scientific models or theories (64, 65).

Ethical challenges also constitute an important area of consideration for ERCs. Data breakdown by class of ethical consideration reveals that the current ethical debate is being largely monopolized by issues of privacy and data protection (Figure 2.3). It was already pointed out, that the ethics of big data should not be reduced to a privacy challenge but it encompasses a number of positive ethical goals (66). Several ethical issues for which Mittelstad and Floridi (22) demanded increased ethical attention still appear largely underexplored. For example, our analysis reveals that issues of data ownership, group-level ethical harms, and the distinction between academic and commercial uses of big data, do not appear as ethical priorities. Furthermore, we observed that issues of fairness and the risk of discrimination compose a relatively small portion of the current ethical spectrum even though the misuse of big data has demonstrably resulted in various forms of ethnic, gender and class discrimination (67). While group-level harms are usually considered outside the purview of ERCs, the dangers of ignoring this type of risk require careful assessment (68). Issues of trust, transparency, accountability, dignity compose an even smaller fraction of the current ethical landscape. We suggest that the ethical review of health-related big data research should explore a broader spectrum of ethical issues. In particular, it should scrutinize more carefully (i) whether and how each project attempts to address the social benefits, if any, of research; (ii) how data subjects involved in the study can exercise control over their data (data control problem); (iii) which measures of accountability are being employed by the

researchers, (iv) whether the collected data can be reused for secondary, including malevolent, purposes (dual use problem) and what measures are implemented to prevent that.

These technical, methodological and ethical challenges should not be seen as sealed rooms. Thematic analysis reveals an intimate interconnection between the three thematic families. For example, the technical problem of data security appears strictly connected to the ethical notion of privacy and the regulatory principle of data protection. Similarly, methodological errors like dataset bias might have detrimental ethical consequences such as racial and gender discrimination. This intimate link between technical and ethical issues highlights the importance of cooperative approaches to study design in big data research through strategies like ethical design of data-collecting technologies, proactive ethical assessment of big data studies and ethical requirement analyses for data-sharing platforms, data storage sites and other digital infrastructures. ERCs should be sensitized to this interconnection and examine how weaknesses in one domain affect other domains of evaluation. Similarly, the interdependence of epistemological and ethical issues, which was already highlighted by Mittelstad and Floridi (22), requires careful consideration by ERCs to prevent that inaccurate study designs or data curation practices result in unintended harms to individuals or groups.

Overall, these findings have three main and direct implications for ERCs. First, the significance and complexity of technical and methodological challenges suggests that members of ERCs should need to acquire stronger technical and methodological expertise to adequately review and evaluate health-related big data studies. This might require specific educational courses or other training activities aimed at strengthening ERC-members' ability to identify technical/methodological problems or inaccuracies, especially those that can result in harms to data subjects or society like data security breaches, database corruption and biased algorithm training. Specialized training modules in data science, bioinformatics and cybersecurity might serve this purpose. In parallel, as emerging from the normative suggestions, ERCs need to consider including experts from the afore listed disciplines within the review board. Since health-related big data is here to stay, new expert profiles are needed during the review process. Data scientists, security experts, bioinformaticians should complement the expertise of clinicians, ethicists and other traditional ERC members. ERC members will need to be equipped with the necessary tools to inspect how the data will be collected, in conformity with which security standards they will be stored and shared, what classification systems will be employed, how

uncertainty will be quantified, what cluster models will be adopted during exploratory data mining etc.

In spite of these important challenges, ERCs might still be faced with uncertainty when reviewing health-related big data studies. Review results indicate that only a tiny fraction of studies (13%) provided specific normative recommendations for ERCs. These are suggestions or proposals for ERCs as to the best course of action. Further thematic analysis reveals a general disagreement and a lack of consensus on what codes of conduct should be prioritized, with some authors (25) favouring the simplification of the ethics review process and others (69) requiring more stringent scrutiny. Nonetheless, five recurring themes could be identified: (i) preventing the dangers of downstream data linkage and inadvertent individual identification; (ii) expanding the purview and involvement of ERCs; (iii) developing a clearer understanding of the risks and benefits of health-related big data research, (iv) harmonizing ethical standards for big data research and (v) rethinking the composition of ERCs. The extremely small fraction of studies providing normative recommendations informed by empirical research (i.e. based on studies involving direct observation or experience such as survey questionnaires or focus groups), further underscores how these recommendations are mostly based on individual viewpoints rather than on solid consensus within the research community.

In the debate on what ERCs should do in relation to health-related big data, the opinion of ERC members is missing. Future empirical research is highly required to explore the needs, views and attitudes of ERC members about health-related big data. Empirical research in this domain could methodologically build upon previous studies involving ethics advisors working in big-data- related areas of research such as genomics governance (70). Combining empirical and normative ethical research in the health-related big data domain would not only benefit the understanding of the current problems that ERCs are facing when reviewing health-related big data studies, but also favour the development of empirically-informed research ethics guidelines (71), hence resulting in better ethical oversight and governance of the health-related big data phenomenon.

Finally, it is legitimate to raise the question of whether ERCs should be the only governance body responsible for the evaluation of biomedical big data research. Given their traditional mandate, which is deeply rooted in the pre-digital era of biomedical research, it might be reasonably argued that ERCs are ill-suited to exercise exclusive ethical oversight on health-related big data research.

Research regulators should consider whether complementary governance mechanisms such as data boards, data security committees or allied bodies are necessary to expand the bandwidth and sensitivity of ethical oversight.

Acknowledgements

This manuscript was supported by the Swiss National Science Foundation under award 407540_167223 (NRP 75 Big Data).

References

1. Lynch C. Big data: How do your data grow? Nature. 2008;455(7209):28. pmid:18769419
2. Stephens ZD, Lee SY, Faghri F, Campbell RH, Zhai C, Efron MJ, et al. Big data: astronomical or genomical? PLoS biology. 2015;13(7):e1002195. pmid:26151137
3. Dumbill E. Making sense of big data. Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA; 2013.
4. Mattmann CA. Computing: A vision for data science. Nature. 2013;493(7433):473–5. Epub 2013/01/25. pmid:23344342.
5. Minelli M, Chambers M, Dhiraj A. Big data, big analytics: emerging business intelligence and analytic trends for today's businesses: John Wiley & Sons; 2012.
6. Kim G-H, Trimi S, Chung J-H. Big-data applications in the government sector. Communications of the ACM. 2014;57(3):78–85.
7. Lyon D. Surveillance, Snowden, and big data: Capacities, consequences, critique. Big Data & Society. 2014;1(2):2053951714541861.
8. Lv Y, Duan Y, Kang W, Li Z, Wang F- Y. Traffic flow prediction with big data: a deep learning approach. IEEE Transactions on Intelligent Transportation Systems. 2015;16(2):865–73.
9. Alyass A, Turcotte M, Meyre D. From big data analysis to personalized medicine for all: challenges and opportunities. BMC medical genomics. 2015;8(1):33.
10. Murdoch TB, Detsky AS. The inevitable application of big data to health care. Jama. 2013;309(13):1351–2. pmid:23549579
11. Andreu-Perez J, Poon CC, Merrifield RD, Wong ST, Yang G-Z. Big data for health. IEEE journal of biomedical and health informatics. 2015;19(4):1193–208. pmid:26173222
12. Ienca M, Wangmo T, Jotterand F, Kressig RW, Elger B. Ethical Design of Intelligent Assistive Technologies for Dementia: A Descriptive Review. Science and engineering ethics. 2017:1–21.
13. Dimitrov DV. Medical Internet of Things and Big Data in Healthcare. Healthcare Informatics Research. 2016;22(3):156–63. PubMed PMID: PMC4981575. pmid:27525156
14. McCabe LL, McCabe ER. Direct-to-consumer genetic testing: access and marketing. Genetics in Medicine. 2004;6(1):58. pmid:14726811 View Article PubMed/NCBI Google Scholar
15. Lovett KM, Mackey TK, Liang BA. Evaluating the evidence: direct-to-consumer screening tests advertised online. Journal of medical screening. 2012;19(3):141–53. pmid:23093732
16. Vayena E, Gasser U. Strictly biomedical? Sketching the ethics of the big data ecosystem in biomedicine. The ethics of biomedical big data: Springer; 2016. p. 17–39.
17. Weber GM, Mandl KD, Kohane IS. Finding the missing link for big biomedical data. JAMA. 2014;311(24):2479–80. pmid:24854141
18. Vayena E, Dzenowagis J, Brownstein JS, Sheikh A. Policy implications of big data in the health sector. Bulletin of the World Health Organization. 2018;96(1):66. pmid:29403102
19. Lee CH, Yoon H-J. Medical big data: promise and challenges. Kidney Research and Clinical Practice. 2017;36(1):3–11. PubMed PMID: PMC5331970. pmid:28392994
20. Binder H, Blettner M. Big Data in Medical Science—a Biostatistical View: Part 21 of a Series on Evaluation of Scientific Publications. Deutsches Ärzteblatt International. 2015;112(9):137. pmid:25797506

21. Bates DW, Saria S, Ohno-Machado L, Shah A, Escobar G. Big data in health care: using analytics to identify and manage high-risk and high-cost patients. Health Affairs. 2014;33(7):1123–31. pmid:25006137
22. Mittelstadt BD, Floridi L. The ethics of big data: Current and foreseeable issues in biomedical contexts. Science and Engineering Ethics. 2016;22(2):303–41. pmid:26002496
23. Vayena E, Salathé M, Madoff LC, Brownstein JS. Ethical challenges of big data in public health. PLoS computational biology. 2015;11(2):e1003904. pmid:25664461
24. Peek N, Holmes J, Sun J. Technical challenges for big data in biomedicine and health: data sources, infrastructure, and analytics. Yearbook of medical informatics. 2014;9(1):42.
25. Balas EA, Vernon M, Magrabi F, Gordon LT, Sexton J, editors. Big Data Clinical Research: Validity, Ethics, and Regulation. MedInfo; 2015.
26. Foster Riley M. Big data, HIPAA and the common rule: time for big change?: Cambridge University Press; 2018.
27. Vayena E, Mastroianni A, Kahn J. Caught in the web: informed consent for online health research. Sci Transl Med. 2013;5(173):173fs6. pmid:23427242
28. Cambridge Analytica controversy must spur researchers to update data ethics. Nature. 2018;555(7698):559–60. Epub 2018/03/30. pmid:29595795.
29. Wang Y, Kosinski M. Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. Journal of personality and social psychology. 2018;114(2):246. pmid:29389215
30. El Emam K, Jonker E, Arbuckle L, Malin B. A Systematic Review of Re-Identification Attacks on Health Data. PLOS ONE. 2011;6(12):e28071. pmid:22164229
31. de Montjoye Y- A, Radaelli L, Singh VK, Pentland AS. Unique in the shopping mall: On the reidentifiability of credit card metadata. Science. 2015;347(6221):536. pmid:25635097
32. Phillips A, Borry P, Shabani M. Research ethics review for the use of anonymized samples and data: A systematic review of normative documents. Accountability in Research. 2017;24(8):483–96. pmid:29140743
33. Kramer AD, Guillory JE, Hancock JT. Experimental evidence of massive-scale emotional contagion through social networks. Proceedings of the National Academy of Sciences. 2014:201320040.
34. Fiske ST, Hauser RM. Protecting human research participants in the age of big data. Proceedings of the National Academy of Sciences. 2014;111(38):13675.
35. Pham MT, Rajić A, Greig JD, Sargeant JM, Papadopoulos A, McEwen SA. A scoping review of scoping reviews: advancing the approach and enhancing the consistency. Research Synthesis Methods. 2014;5(4):371–85. PubMed PMID: PMC4491356. pmid:26052958
36. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. International journal of social research methodology. 2005;8(1):19–32.
37. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implementation science. 2010;5(1):69.
38. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative research in psychology. 2006;3(2):77–101.
39. Gentles SJ, Lokker C, McKibbon KA. Health information technology to facilitate communication involving health care providers, caregivers, and pediatric patients: a scoping review. Journal of medical Internet research. 2010;12(2):e22. Epub 2010/06/22. pmid:20562092; PubMed Central PMCID: PMCPMC2956233.

40. Peterson J, Pearce PF, Ferguson LA, Langford CA. Understanding scoping reviews: Definition, purpose, and process. Journal of the American Association of Nurse Practitioners. 2017;29(1):12–6. pmid:27245885
41. Cameron JI, Tsoi C, Marsella A. Optimizing stroke systems of care by enhancing transitions across care environments. Stroke. 2008;39(9):2637–43. Epub 2008/07/19. pmid:18635864.
42. Levac D, Wishart L, Missiuna C, Wright V. The application of motor learning strategies within functionally based interventions for children with neuromotor conditions. Pediatric physical therapy: the official publication of the Section on Pediatrics of the American Physical Therapy Association. 2009;21(4):345–55. Epub 2009/11/20. pmid:19923975.
43. Luo J, Wu M, Gopukumar D, Zhao Y. Big data application in biomedical research and health care: a literature review. Biomedical informatics insights. 2016;8:BII. S31559.
44. Research S. Big Data in the Healthcare & Pharmaceutical Industry: 2017–2030 – Opportunities, Challenges, Strategies & Forecasts. 2017:1–499.
45. Signorini A, Segre AM, Polgreen PM. The use of Twitter to track levels of disease activity and public concern in the US during the influenza A H1N1 pandemic. PloS one. 2011;6(5):e19467. pmid:21573238
46. Robinson J, Cox G, Bailey E, Hetrick S, Rodrigues M, Fisher S, et al. Social media and suicide prevention: a systematic review. Early intervention in psychiatry. 2016;10(2):103–21. pmid:25702826
47. Ito K, Weinberger KR, Robinson GS, Sheffield PE, Lall R, Mathes R, et al. The associations between daily spring pollen counts, over-the-counter allergy medication sales, and asthma syndrome emergency department visits in New York City, 2002–2012. Environmental Health. 2015;14:71. PubMed PMID: PMC4549916. pmid:26310854
48. DellaValle CT, Triche EW, Leaderer BP, Bell ML. Effects of Ambient Pollen Concentrations on Frequency and Severity of Asthma Symptoms Among Asthmatic Children. Epidemiology (Cambridge, Mass). 2012;23(1):55–63. PubMed PMID: PMC3246281. pmid:22082997
49. DeMartino JK, Larsen JK. Data needs in oncology: "Making sense of the big data soup". JNCCN Journal of the National Comprehensive Cancer Network. 2013;11(SUPPL2):S1–S12.
50. Benedict SH, El Naqa I, Klein EE. Introduction to Big Data in Radiation Oncology: Exploring Opportunities for Research, Quality Assessment, and Clinical Care. International Journal of Radiation Oncology Biology Physics. 2016;95(3):871–2. PubMed PMID: WOS:000377370600002. pmid:27302502
51. Kansagra AP, John-Paul JY, Chatterjee AR, Lenchik L, Chow DS, Prater AB, et al. Big data and the future of radiology informatics. Academic radiology. 2016;23(1):30–42. pmid:26683510
52. Nekrutenko A, Taylor J. Next-generation sequencing data interpretation: enhancing reproducibility and accessibility. Nature Reviews Genetics. 2012;13(9):667. pmid:22898652
53. Monteith S, Glenn T, Geddes J, Whybrow PC, Bauer M. Big data for bipolar disorder. International Journal of Bipolar Disorders. 2016;4(1). pmid:27068058
54. Christen M, Biller-Andorno N, Bringedal B, Grimes K, Savulescu J, Walter H. Ethical challenges of simulation-driven big neuroscience. AJOB Neuroscience. 2016;7(1):5–17.
55. Vie LL, Griffith KN, Scheier LM, Lester PB, Seligman ME. The Person-Event Data Environment: leveraging big data for studies of psychological strengths in soldiers. Front Psychol. 2013;4:934. Epub 2014/01/01. pmid:24379795; PubMed Central PMCID: PMCPmc3861613.

56. Roccetti M, Prandi C, Salomoni P, Marfia G. Unleashing the true potential of social networks: confirming infliximab medical trials through Facebook posts. Network Modeling and Analysis in Health Informatics and Bioinformatics. 2016;5(1).
57. Megherbi DB, Soper B, editors. Analysis of how the choice of machine learning algorithms affects the prediction of a clinical outcome prior to minimally invasive treatments for benign pro static hyperplasia BPH. CIMSA 2012–2012 IEEE International Conference on Computational Intelligence for Measurement Systems and Applications, Proceedings; 2012.
58. Anderson SL, Anderson M, editors. A Prima Facie duty approach to machine ethics and its application to elder care. AAAI Workshop—Technical Report; 2011.
59. Azaria A, Ekblaw A, Vieira T, Lippman A, editors. Medrec: Using blockchain for medical data access and permission management. Open and Big Data (OBD), International Conference on; 2016: IEEE.
60. Krawiec R, Housman D, White M, Filipova M, Quarre F, Barr D, et al., editors. Blockchain: Opportunities for health care. Proc NIST Workshop Blockchain Healthcare; 2016.
61. Vayena E, Blasimme A. Health Research with Big Data: Time for Systemic Oversight. The Journal of Law, Medicine & Ethics. 2018;46(1):119–29.
62. Popper K. The logic of scientific discovery: Routledge; 2005.
63. Mayer-Schönberger V, Cukier K. Big Data: A Revolution That Will Transform How We Live, Work, and Think. Citeseer; 2014.
64. Mazzocchi F. Could Big Data be the end of theory in science?: A few remarks on the epistemology of data-driven science. EMBO reports. 2015;16(10):1250–5. pmid:26358953
65. Geerts H, Dacks PA, Devanarayan V, Haas M, Khachaturian ZS, Gordon MF, et al. Big data to smart data in Alzheimer's disease: The brain health modeling initiative to foster actionable knowledge. Alzheimer's & dementia: the journal of the Alzheimer's Association. 2016;12(9):1014–21.
66. Vayena E, Tasioulas J. The dynamics of big data and human rights: the case of scientific research. Phil Trans R Soc A. 2016;374(2083):20160129. pmid:28336802
67. Beck EJ, Gill W, De Lay PR. Protecting the confidentiality and security of personal health information in low- and middle-income countries in the era of SDGs and Big Data. Glob Health Action. 2016;9:32089. Epub 2016/11/26. pmid:27885972; PubMed Central PMCID: PMCPmc5123209.
68. Mittelstadt BD, Floridi L. The Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts. Science and engineering ethics. 2016;22(2):303–41. Epub 2015/05/24. pmid:26002496.
69. Rothstein MA. Ethical Issues in Big Data Health Research: Currents in Contemporary Bioethics. The Journal of Law, Medicine & Ethics. 2015;43(2):425–9. pmid:26242964
70. Samuel GN, Farsides B. Public trust and 'ethics review'as a commodity: the case of Genomics England Limited and the UK's 100,000 genomes project. Medicine, Health Care and Philosophy. 2017:1–10.
71. Zook M, Barocas S, Crawford K, Keller E, Gangadharan SP, Goodman A, et al. Ten simple rules for responsible big data research. PLoS computational biology. 2017;13(3):e1005399. pmid:28358831

CHAPTER 3: Machine Learning in Medicine: Opening the New Data Protection Black Box

Peer reviewed and published as: Ferretti, A.*, Schneider, M.*, Blasimme, A. (2018). Machine Learning in Medicine. *European Data Protection Law Review*, 4(3), doi:10.21552/edpl/2018/3/10.

* These authors contributed equally and share first authorship

3.1 Abstract

Artificial intelligence (AI) systems, especially those employing machine learning methods, are often considered black boxes, that is, systems whose inner workings and decisional logics remain fundamentally opaque to human understanding. In this article, we set out to clarify what the new General Data Protection Regulation (GDPR) says on profiling and automated decision-making employing opaque systems. More specifically, we focus on the application of such systems in the domain of healthcare. We conducted a conceptual analysis of the notion of opacity (black box) using concrete examples of existing or envisaged medical applications. Our analysis distinguishes among three forms of opacity: (i) lack of disclosure, (ii) epistemic opacity, and (iii) explanatory opacity. For each type of opacity, we discuss where it originates from, and how it can be dealt with according to the GDPR in the context of healthcare. This analysis can offer insights regarding the contested issue of the explainability of AI systems in medicine, and its potential effects on the patient-doctor relationship. Keywords: Artificial Intelligence, Machine Learning, Black Box, Medicine, GDPR, Transparency

3.2 Introduction

Artificial intelligence (1) (AI) is the talk of the town. In recent years, we have witnessed a growing interest for the rapid development and application of AI systems in virtually any domain of human activity. Even objects of ordinary use such as thermostats, smartphones and cars now employ AI systems to process data and automatically perform an increasing number of tasks. In particular, some fields of AI such as machine learning (ML) are attracting attention in both academic circles and the popular press. These systems have the ability to learn, on their own or through human supervision, how to perform a task such as, for instance, recognizing a road sign. But AI systems are also employed to automate certain judicial decisions, for instance to help judges predict the risk of recidivism (2), or in the financial industry to aid decisions about loans (3) or insurance policies (4).

Another very relevant area of application is medicine. In 2016, a paper in JAMA by Gulshan and colleagues showed that an AI system was able to identify diabetic retinopathy and diabetic macular oedema in retinal fundus images with a degree of accuracy similar to that of licensed

ophthalmologists (5). One year later, on the same journal, a research group form the Netherlands demonstrated that an AI system employing deep learning (a form of machine learning) is able to detect nodal metastases in women with breast cancer with the same accuracy as clinical pathologists (6). A medical AI system that interprets magnetic resonance images in cardiology and radiology has already been licensed for use (7). Moreover, earlier this year, the FDA gave clearance to an AI system that helps orthopaedists detect wrist fractures in two-dimensional x-ray scans (8).

While most people recognize the promise of applying AI systems to medical diagnosis and decision-making, many are worried about the use of partly autonomous computer programs for medical purposes. This fear has to do with a characteristic of many ML methods. AI systems that incorporate ML learn with a varying degree of supervision which rules (9) they need to follow in order to perform their task. The programmer sets up the system so that it can learn to do something. However, he or she does not decide, nor is necessarily aware of the rules the AI system has learnt and is following in order to do what it is supposed to do. This characteristic is often referred to as the opacity of ML. For the same reason, AI systems based on ML are often called black boxes, to stress that it is hard or even impossible for human users to open them up, so to say, and see for themselves what the machine is doing (or, which is the same, what rule the machine has learnt and is employing). The possibility that these systems could remain opaque to their own creators as well as to their end-users is a cause of concern (10).

The issue of opacity in AI systems for medical applications is only starting to be discussed. For example, Char and colleagues have recently argued that pressing ethical challenges loom large on the horizon of ML in healthcare, precisely because of opacity (11). In particular, they point out that physicians lack adequate education to understand the construction and limitations of such systems and they stress that: ‘Remaining ignorant about the construction of machine-learning systems or allowing them to be constructed as black boxes could lead to ethically problematic outcomes’ (12).

The debate about lack of algorithmic transparency has taken several directions. Some scholars have argued for a presumed ‘right to explanation’ for data subjects whose data is processed by means of AI systems. According to this alleged right, when profiling or automated decisions affect peoples’ capacity to access certain goods and services, data subjects have a right to be provided with adequate explanations regarding the processes that led to those outcomes. The idea of a

right to explanation stems from the value of transparency in data processing and it is intended to counterbalance the opacity of automated systems. However, its actual definition and scope are contested. Some of its proponents – like Wachter, Mittelstadt and Floridi – distinguish between the explanation of a system's general functionality, and the explanation of specific decisions taken through or by an artificial intelligence system (13). According to this group of authors, explaining an automated decision-making system's functionality means explaining its 'logic, significance, envisaged consequences, and general functionality' (14). On the other hand, explaining specific decisions means explaining 'the rationale, reasons, and individual circumstances of a specific automated decision, eg the weighting of features, [or] machine-defined case-specific decision rules' (15). According to Wachter and colleagues, only the latter form of ex post explanation of specific decisions genuinely fulfils the idea of a right to explanation. This understanding of the right to explanation has been fiercely criticised. Selbst and Powels, for instance, have argued that such framework overlooks the fundamental characteristics of AI systems (16).

Interestingly, this debate has been triggered by some provisions already present in the European Data Protection Directive (17) (DPD) and now restated and expanded in the European Data Protection Regulation (18) (GDPR). The GDPR establishes principles, obligations and rights in the context of profiling and automated individual decision-making. Some of these new provisions entitle data subjects to receive information, explanation and protection (in the forms of rights and safeguards) regarding profiling, automated decisions, and special categories of data involved in such activities. In particular, data subjects are entitled to receive meaningful information about the logic involved, the significance and the envisaged consequences of solely automated individual decision-making and profiling activities as stated in Articles 13(2)(f), 14(2)(g) and 15(1)(h) of the GDPR.

Some scholars – including Selbst and Powels – have recognized in these provisions the implementation of a right to explanation (19). Others instead – like Wachter and colleagues – do not think that existing provisions in the GDPR adequately address the full scope of this right (20). It is certainly true that the GDPR mentions 'explanation' only once, in Recital 71 – and that recitals are non-binding. Yet, while the GDPR does not explicitly refer to the issue of opacity, nor to the metaphor of the black-box, it nonetheless specifies a set of rights, safeguards and conditions that require data controllers to communicate some relevant features of profiling and automated data-processing systems to data subjects (21). These provisions, apply also to AI systems in the domain of healthcare, and therefore they oblige data controllers (in this case, hospitals and

physicians) to provide meaningful information to patients about the use of such systems. Yet, the implementation of such provisions in the clinical setting calls for a deeper analysis of what opacity of AI systems amounts to and how data controllers can fulfill the demands for greater transparency towards a specific category of data subjects – a category that, in this case, will likely correspond to patients. This analysis is intended to dispel some of the potential sources of confusion that may have affected current debates on the right to explanation. This will in turn facilitate the legal interpretation of the GDPR provisions on profiling and automated decision-making in the domain of healthcare, hopefully fostering consensus on efficient and yet meaningful ways to fulfil data controllers' obligations in this rapidly evolving area of medical technology.

In what follows we will first summarise the GDPR provisions in relation to profiling and automated decision-making (II), focusing on rights and obligations regarding communication to data subjects about such activities. We then provide an analysis of the notion of opacity, its sources and how the GDPR provisions could be fulfilled in the clinical setting (III). We then discuss broader ethical and practical implications of such obligations in the context of the patient-doctor relationship (IV).

3.3 Promoting Transparency: What Does the GDPR Demand

The new GDPR is driven by the general aim of improving the protection of data subjects (22). A further aim of the Regulation is that of enhancing overall transparency of data processing activities (23). The principle of transparency is defined as requiring 'that any information and communication relating to the processing of … personal data be easily accessible and easy to understand, and that clear and plain language be used' (24). In particular, data subjects should be informed about the identity of data controllers and the aims of data processing, they should obtain confirmation and communication of data processing activities employing their data and should be made aware of their rights. Further safeguards should be in place to limit data collection and use only the strict minimum amount of data, to ensure privacy and to offer the possibility of rectifying or deleting data (25). These general principles apply to all kinds of data processing activities, and underlie a number of rights, obligations and requirements (26). However, special provisions exist that define rights and obligation in the case of profiling and automated decision-making. Those parts of the GDPR are of direct relevance to our present analysis, since AI systems are most likely going to fall within this type of data processing activities. The GDPR defines profiling as

any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interest, reliability, behavior, location or movement (27).

Automated decision-making can take place by means of profiling or not; similarly, profiling can happen with or without automated decision-making. Moreover, automated decision-making can be made with or without human involvement – in the latter case it amounts to what the GDPR defines as solely automated decision-making.

The use of AI systems for health-related purposes clearly falls under these definitions, and therefore counts as a form of profiling within the GDPR (28). This is confirmed by the Data Protection Working Party Guidelines on automated decision-making and profiling explaining the GDPR provisions in this specific domain (29). The guidelines explicitly state that profiling is generally employed 'to make predictions about people, using data from various sources to infer something about an individual' (30). Moreover, the Guidelines note that profiling and automated decision-making can pose significant risk to the data subjects since 'these processes can be opaque' (31). This reference to that highly debated characteristic of AI systems indicates the attention of the legislators to the effects of opacity on the rights and interests of data subjects.

The specific focus on profiling and automated decision-making in the GDPR gives rise to a rather complex set of provisions that we will now briefly illustrate. We focus mainly on provisions prescribing communication and explanations of profiling (32), decisions made by a human based on profiling (33), and solely automated decision-making (34).

As it is customary, we will start this overview from the indications contained in non-binding recitals (35). Recital 63 states that 'every data subject should … have the right to know … the logic involved in any automatic personal data processing and, at least when based on profiling, the consequences of such processing' (36). Recital 71, in a similar vein, stresses that legally authorised profiling and solely automated decisions can only take place if data subjects 'right … to obtain an explanation of the decision reached after such assessment' is respected (37).

These indications clearly recall the need to counterbalance the opacity of certain forms of data processing like AI systems and to promote transparency around activities such as profiling,

decision-making based on profiling and solely automated decision-making producing legal effects or similarly significant consequences for data subjects (38).

As far as binding provisions are concerned, Article 5(1)(a) establishes transparency as a basic principle of data processing (along with lawfulness and fairness). As explained by the Guidelines, this principle is especially relevant to profiling because this type of activity is often invisible to data subjects, but also because data subjects may not have the necessary capacities to understand the technical aspects of profiling and automated decision-making activities – which represents, as we will see below, two distinct ways of articulating opacity (39).

The GDPR recognizes specific rights to data subjects that derive from the general principle of transparency. They apply to all data processing activities and articulate precise requirements concerning what data controllers are supposed to communicate to data subjects regarding the activities they intend to conduct or are conducting with their data (40).

Some communication-related requirements apply exclusively to solely automated decision-making and profiling activities. Overall, data subjects have a 'right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her' (41). This type of activity is in principle prohibited by Article 22(1) but exceptions apply as per Article 22(2): (a) when the processing is needed to enter or perform a contract between data subjects and data controllers; (b) when it is authorised by the Union or a Member State; and (c) when data subjects express explicit consent.

When special categories of data are employed in solely automated decision-making or profiling, exceptions (a) and (c) do not apply, unless the data subjects have explicitly consented to the use of such data, or processing is justified by a substantial public interest (42). Under special categories, the GDPR lists personal data

revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data … , data concerning health or data concerning a natural person's sex life or sexual orientation (43).

On the basis of this definition, solely automated decision-making and profiling in the context of medicine by means of AI systems amount to activities employing special categories of data.

As far as communication to data subjects is concerned, a data controller should inform them of the existence of solely automatic decision-making and profiling activities, and should provide 'meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject' (44). The same applies to activities based on data not obtained directly from the data subject (45). Data subjects also have a right to access all their personal data used in the context of such activities, as well as all information relative to profiling itself, that is, the categories used to set up the profiling and the classifications applied to the data subject as a result of the profiling activity (46).

The communication of the 'logic involved, [and of] the significance' of solely automated data processing brings up the issue of explainability (47). What does it mean to inform a data subject about the logic and significance of an AI system, especially using ML, in the context of healthcare? Does opacity stand in the way of fulfilling these requirements? And if so, in which specific way? To address these questions, in the next section we offer a conceptual analysis of the notion of opacity with the aim of clarifying its implications in the contest of the GDPR.

3.4 Unpacking the Notion of Opacity

The idea that AI systems can be opaque black boxes, and that this characteristic poses risks to humans who are affected by their decisions has recently gained traction. Frank Pasquale, in his 2015 book The Black Box Society, was among the first to popularise the idea that AI systems are 'black-boxes' (48). According to Pasquale, AI systems deserve this label because they pervasively collect and compute personal data to help companies pursue their business aims, while most data subjects are unaware such activities are even taking place. This may be the case because firms intentionally conceal their profiling activities in order to avoid objections on the part of data subjects or because they want to prevent competitors from becoming aware of their trade secrets (49). This view articulates a first, very common understanding of opacity, that we found also in the Guidelines to the GDPR stating that 'individuals might not know that they are being profiled' (50) and that they may not know that data not provided by themselves can also be used to profile them (51). We call this form of opacity 'lack of disclosure'.

The notion of opacity however possesses also other semantic connotations. For instance, the image of a black box has been extended to the application of AI systems in medicine (52). Price describes what he calls 'black-box medicine' as a form of medicine heavily based on algorithms

that can analyse large amounts of data. Black box medicine relies on opaque algorithms to find patterns among patient data that can, for instance, make predictions about health-related risks, or suggest a better dosage of a drug. For example, one can imagine the case of a woman living in a specific urban environment and presenting a specific combination of genetic variants. A ML-based AI system may process such data and predict that the patient has, say, a higher-than-average risk of developing aggressive breast cancer. Price argues that such an algorithm can be opaque to the data subject, as well as to the physician, in two respects. On the one hand, the ML system may discover patterns within such a high number of variables that it is extremely hard for a human mind to make sense of. This type of opacity is the result of the sheer complexity of data computation – relying in our case on millions of genetic and environmental factors interacting with each other. We can refer to it as 'epistemic opacity' – to stress that opacity here originates from a lack of understanding regarding how the ML system operates. On the other hand, opacity may amount to the fact that, while an ML system can identify patterns in an incredibly high number of variables, it may not be possible to trace them back to any known causal explanation of the association between input and output, that is, between patient data and the prediction made by the system. In other words, since the ML system is not programmed with any particular clinical hypothesis in mind (eg the hypothesis that certain genetic mutations increase the risk of breast cancer), it may discover patterns that, at the present stage of medical knowledge, cannot be linked to any known causal explanation. We can call this kind of opacity 'explanatory opacity', to stress that it refers to the lack of a clinical explanation or, that is the same thing, to the impossibility of clinically interpreting the outcome of the ML system.

This rapid excursus shows that there are at least three different ways of interpreting the meaning of opacity: as lack of disclosure, as epistemic opacity, or as explanatory opacity (53). These three semantic dimensions of opacity are particularly relevant to the application of AI systems in medicine. However, this conceptual framework can be fruitfully applied to the any other use of AI systems that fall under the GDPR provisions.

In what follows we will discuss each of these three semantic dimensions of opacity in the context of the requirements, rights and obligations stipulated by the GDPR in relation to automated decision-making and profiling activities.

Lack of Disclosure

From a general point of view, lack of disclosure refers to the fact that data subjects are unaware that automated decision-making and profiling activities about them are being carried out. This type of opacity does not depend on intrinsic technical characteristics of AI systems, but derives from the way automated data processing and profiling activities are conducted. In principle, any type of data processing activity could be conducted without the data subjects being aware of it. It follows that lack of disclosure is not specific to the use of AI systems. Yet in the field of automated decisions and profiling, this kind of opacity prevents data subjects from exercising some specific data-related rights, such as the right to object to solely automated data processing as stated in Article 13(2)(f). This may have very tangible consequences in the context of medical applications.

Lack of disclosure may depend on intentional concealment of relevant information regarding automated data processing. This may be linked to the attempt on the part of data controller to avoid interference with or potential objections to their activities, but it can also be due to the need to protect intellectual property, copyright and trade secrets. Since algorithms and data are non-rival goods, once disclosed, their value can be dramatically reduced unless protected through patents. Therefore, there is lack of transparency when AI systems are subject to non-disclosure policies. Particularly, private corporations might try to use trade secrets to protect against competitors. Recital 63, although not legally enforceable, affirms that data subjects should have the right to know 'the logic involved in any automatic personal data processing and, at least when based on profiling, the consequences of such processing' (54). This information shall be disclosed to data subjects pursuant Articles 13 and 14 on information, and Article 15 on data access. Yet, the same Recital 63 recognizes that such disclosure shall not adversely affect trade secrets, intellectual property and copyrights on software. It follows that, while data controllers must disclose that they are conducting profiling or automated data processing, they are not obliged to reveal all details about their AI systems. In practical terms, this entails that data controllers may still be required to provide information regarding the general characteristics of their system, but they may not be compelled to explain what rules the AI system follows, how it has reached a conclusion, or how it has taken a given decision about a particular data subject.

Lack of disclosure may also depend on the fact that the data controller is employing data not directly provided by data subjects, but inferred from other data available to the data controller. In this case, the data subject will most likely remain unaware of profiling activities about him or herself unless the data controller communicates details about such activities. Now, according to

the GDPR, data subjects have a right to receive a number of information regarding who is processing their data, for which purposes and under which circumstances (55). Yet data controllers are not obliged to disclose the technical nature of their data processing activities. If, however, data processing amounts to solely automated processing (56) producing legal effects or similarly significant consequences, the Regulation stipulates that they should inform data subjects about or confirm the existence of such form of processing, and provide 'meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject' (57). This applies also to uses of data that have not been directly obtained from data subjects (58). This type of information should happen already at the moment of data collection (59). However, such provisions do not apply to non-solely automated profiling as defined in Article 4 (60). It follows that data controllers are not obliged to communicate the existence, the logic and the foreseeable consequences of their automated decision-making and profiling activities if they include even a minimal degree of human involvement.

In the medical context, this form of disclosure appears all the more important given the rapid growth of AI-based automated decision-making tools in the practice of medicine. In this case, disclosure other than being legally mandatory, is also justified by the need to promote patients' trust in the use of AI systems in healthcare. Failing to reveal to the patient that such systems are in use may undermine the fiduciary relationship between patients and doctors, and may also give patients the impression that they are being marginalized in decisional processes regarding their health, thus affecting their decisional autonomy and their sense of self-determination. In light of these considerations, restricting disclosure to solely-automated activities may turn out to be insufficient (61). Informed consent seems to be the most obvious way to discharge such informational obligations regarding the clinical use of AI systems (62). However, research is needed to show whether conventional informed consent procedures do actually constitute a valid means to convey this type of information and to increase patients' trust. Other means of communication, including innovative forms of consent (63) may actually prove equally or more effective than conventional written consent.

Epistemic Opacity

Epistemic opacity occurs when it is not possible to have access to or there is not sufficient understanding of the rules an AI system is applying to make predictions, classifications and

decisions. Epistemic opacity is therefore related to the question of how an AI system provides a specific outcome.

This type of opacity can have two sources: a) procedural darkness or b) procedural ignorance.

Procedural Darkness

Procedural darkness occurs when the developers or the users themselves (and as a consequence, data subjects) do not have access to the rules that the AI system is following to produce a certain output. This is a consequence of how ML systems work. A ML system can learn the rules that it will apply to make classifications and decisions based on given input data. This means that neither the developer, nor the user (eg a physician or a healthcare provider) is aware of how the system produces its outputs. This information is usually encoded in an abstract form in the parameters that the ML system has learnt and therefore it is not readily accessible in a semantically readable form. As a consequence, programmers and users may not be in a position to disclose meaningful information about the way a given output has been reached. While this does not prevent data controllers from explaining the general working principles of the systems, it limits their capacity to reveal exactly which features of the input data have been taken into account to arrive at a given decision about a given data subject. This does not mean, however, that such explanation is impossible, and that this type of opacity cannot be dispelled – and actually the GDPR does not require this type of explanation. Moreover, it is unclear whether the patients' rights and interests are truly promoted by getting access to this type of information.

Procedural Ignorance

Even assuming that the rules of an AI system are accessible in a semantic form, acquiring a meaningful understanding of their role in data processing activities may require a considerable amount of background knowledge. In the case of AI systems used in oncology for instance, a patient should possess basic notions of both cellular pathology and computer science to make sense of what an AI system for automated pathology screening does when it looks at scanned images of human biopsies. In the presence of procedural ignorance, the patient cannot grasp information about the rules the AI system is following. This can also happen because such rules are simply too many. The same problem can occur with attempts at explaining which rules the system applies in his or her specific case. Moreover, it cannot be assumed that physicians

possess enough insight into these matters either. Patient education sessions and specifically trained consultants could be helpful to dispel procedural ignorance. In general, however, it remains to be seen whether dispelling procedural ignorance is indeed needed to ensure patients meaningfully consent to the use of AI systems in healthcare. The GDPR seems more oriented to explanations that do not dwell into the technical details of automated data processing and profiling activities, but rather provide accessible yet meaningful explanation of the general principles involved in their design (64). The aim of such provision, seems linked to the possibility of exerting the rights that the Regulation recognizes to data subjects, including the right to object to fully automated data processing and profiling.

Epistemic opacity limits the possibility of providing thorough explanations of either the inner workings of AI systems or the specific rules used to make a specific, individual decision. However, these limitations do not entail the impossibility of explaining the foreseeable consequences of such processing (65). Moreover, according to the Guidelines, data subjects who are the object of profiling activities are still entitled to access the 'details of any personal data used for profiling, including the categories of data used to construct a profile … and details of which segments the data subject has been placed into' (66).

AI systems can operate on a too complex level of inputs to be understandable by humans. For example, machine learning is able to predict whether a patient has diabetes (67) or lung cancer (68) including hundreds of thousands of heterogeneous data (eg from images, personal health records, lifestyle data collected through wearables devices and so on) in multiple combinations (69). Medical algorithms can also make predictions about the response of a tumour to a specific drug looking at allelic patterns among thousands of genes (70). Yet, prognostic models commonly used by physicians until recently were restricted only to a relatively limited number of variables (71). Even in the case in which the physician could in principle understand the datasets and the rules used by the machine, complexity may still hinder his or her full understanding of how the system operates. The information available to the physician would be either too large in volume or computed in such an elaborated way to be nearly impossible to grasp at a cognitive level. Neither the doctor nor an expert of AI systems might be able to understand in detail how the data is processed exactly and how the output is computed from the data (72).

It should be noted, however, that AI specialists are trying to reduce procedural darkness by technical means. Different techniques allow deeper insights into AI systems. In image recognition,

for example, a method called deconvolution provides information about what rules very complex systems like artificial neural networks learn (73). Deconvolution allows to visualise otherwise internal – and inaccessible – states of the system but the rules the system uses for its decision-making have to be derived experimentally. Such a method could lead researchers to realize that, for instance, some particular properties of an image (eg shapes, edges, colours) are what the system is looking at to determine the outcome. While this is a very interesting area of research, it remains to be seen whether and how such knowledge could one day be employed to provide more detailed explanations to data subjects. In the case of clinical applications, then, these considerations will also need to take into account the specific clinical circumstances at stake, as well as the best interests of the patient.

Explanatory Opacity

Explanatory opacity relates to the question of why an AI system provides a specific outcome. What an ML system is designed to do is to discover patterns between huge numbers of variables in a training dataset, and to leverage these patterns to make classifications, predictions and decisions regarding new input data. The output, in other words, is the result of patterns that the ML system has generalised from training examples. In the field of medicine, for example, an ML system could learn that certain geometrical properties of a histology slide correlate with a bad prognosis. In an extremely simplified scenario, we can imagine that the system learns that a given cellular shape is linked to a bad prognosis. It may well be that this rule corresponds exactly to one of the criteria that pathologists use to recognize an aggressive tumour – and therefore to adequately predict a bad prognosis. Pathologists may also have a scientific explanation of the reason why aggressive tumour cells acquire that shape and behave the way they do. It could be, however, that the ML system looks at properties like relative pixel luminosity that human pathologists do not use, and whose connection to a given clinical phenotype is not known. In those cases, one could confidently say that there is a statistically relevant correlation between the property and the clinical classification. Still, one would not have a scientific explanation of the reason why the property and the classification are linked one to the other (74). In order to possess such an explanation, one should know the biological mechanism that connects the observed property with the predicted phenotype, or, in other words, one should know a causally relevant

chain of biological facts underlying the association between the two (75). We have called the absence of this kind of knowledge 'explanatory opacity'.

This characteristic is typical but not exclusive of knowledge produced through ML systems. There are in fact many examples in medicine of perfectly reliable correlations that lack a scientific explanation. For instance, it has long been known that being heterozygous for a point mutation in the β-chain of the human haemoglobin gene (a mutation that in homozygosity causes sickle cell anaemia) provides protection against malaria (76). Yet only recently has the biological mechanism by which this mutation confers protection been elucidated. Analogously, very little is known (77) about the reason why acetaminophen (paracetamol) works, and yet this is one of the most common pain killers and antipyretic drugs worldwide. Examples of other drugs currently in use despite limited knowledge regarding why they work are countless (78). This shows that explanatory opacity does not represent a dead end in the context of medicine. Little research exists however to analyse whether opaque AI systems in medicine pose specific risks to patients. Also, it is not clear whether the notion of 'meaningful information' to be communicated to data subjects in the case of automated data processing and profiling may include providing causal mechanistic explanations of algorithmic decisions. The Guidelines suggest that this is not the case, and that explanations should focus on the rationale for data processing rather than on its scientific interpretability (79).

It could be argued, however, that reduced explanatory opacity constitutes a reassurance regarding the clinical validity of medical AI systems. Moreover, physicians may be more likely to adopt such systems if they could be provided with clinically intelligible reasons for using them. Yet, this does not imply that this type of information should be communicated to patients, nor that they have a right to it. It may be – and more research could help find out – that patients are not interested in this type of detailed scientific explanations. Moreover, it is still not known whether this type of explanations have an impact on patients' trust in the use of these technologies. What is certainly relevant for promoting the rights and interests of patients is that AI systems used in healthcare respond to the highest standards of safety and efficacy. Careful clinical validation processes by regulatory agencies and through self-imposed industry standards will have to be adopted, and the implications of explanatory opacity should be duly taken into account in the development of such standards.

An alternative to providing in-depth information about an AI system and a way to overcome the implications of explanatory opacity lies in counter-factually examining the outcomes provided by the AI system. Wachter, Mittelstadt and Russell have argued that 'counterfactuals describe a dependency on the external facts that led to that decision' and that more than one counterfactual might exist (80). Counterfactuals can illustrate how even a small change in an input variable can result in a different outcome. Input variables can therefore be systematically varied and corresponding outcomes can be compared to the original, in order to gain deeper insight into the relationship between input variables and the computed outcomes. This opens up the possibility to infer a reason why the original outcome was produced. The explanation provided based on a specific case might not result in an exhaustive description of the entire AI system, but it would contribute significantly to the perspective of human understanding (81).

3.5 Discussion

With the rapid development of AI systems for medical use, patient data will increasingly be used in automated data processing, profiling and decision-making activities. These activities are not restricted to data types that clearly fall within the remit of health data. As a matter of fact, an increasing variety of data generated and collected outside the clinical setting, and not initially intended for medical use are now starting to be employed in diagnosis, health-risk predictive models and to guide medical decisions (82). These include, for instance, lifestyle data, data about dietary habits, socio-economic data, but also data such as keystroke dynamics (83), and in general data collected through smartphones (84) or wearable devices (85). In this expanding health data ecosystem (86), new medical paradigms such as precision medicine (87) and digital health (88) are rapidly growing with the aim of exploiting novel capabilities in data mining and automated data processing for the benefit of patients. These developments go hand in hand with the increasingly perceived need to promote data sharing so as to accelerate the pace of progress in this domain (89). Data sharing, however, needs to be balanced against the risks that unauthorised access and misuse of personal data may pose to data subjects. The GDPR sets out to mitigate such risks regarding pervasive data processing. While health-relevant data are not the main focus of this law, its provisions clearly apply also to the use of data in the medical field. Patients therefore have additional rights as data subjects. Physicians and healthcare providers in general have specific responsibility as data controllers. Such entitlements and obligation are likely to acquire increasing importance with the rapid development and deployment of AI systems in the practice of medicine. The opacity of these systems, however, calls into question their reliability

and trustworthiness. If left unattended, this issue may undermine the development of these technologies and forego the much-awaited benefit they promise to deliver to patients.

Our analysis has shown that opacity is a polysemic concept. This notion is used by different people to indicate different practical and technical characteristics of AI systems. We also showed that not all dimensions of opacity are equally amenable to be dispelled by offering information and explanations to data subjects. However, the GDPR requirements about opacity seem not to require that data subjects are provided with extensive and detailed technical information about the working logic of AI systems processing their data. At least in this respect, the GDPR adopts a not too expansive understanding of transparency.

In the specific domain of healthcare, the amount and type of information provided to patients has long been a topic of discussion due to its practical and ethical implications (90). More research is needed to understand patients' and physicians' attitudes towards opacity in AI systems. The explainability of such systems and the obligations relative to the communication of information about them are likely to affect the patient-doctor relationship. Solely automated decisions and profiling can de-personalize such relationship, reducing occasions for direct personal interaction. Patients may feel undermined from the point of view of their decisional autonomy and capacity to influence healthcare provision practices about themselves. Reasonable communication could mitigate this effect, be a precondition for objecting to fully automated processing, and for requiring some form of human involvement in those activities. This would foster ideals of shared decision making in medicine (91). Also, evidence is needed to understand if written informed consent is the most effective way of communicating information in this domain, or whether other forms of communication could be fruitfully explored.

Another ethically relevant aspect of opaque AI systems is that they may provide predictive information that is not actionable or difficult to interpret. If AI systems create the basis for the unmediated provision of such medical information directly to patients, the information burden on patients may be disproportionately big. Ad hoc safeguards need to be in place to ensure that professional figures stay in the loop of this novel ways of producing medical information and providing healthcare. This raises a thorny issue that, to date, has not received sufficient attention. Paradoxically, according to the GDPR, if a human actor is involved in the operations of an AI system, the obligation to disclose automated data processing and to provide meaningful information about its logic, significance and consequences no longer holds (92). Therefore, it is

precisely the presence of a potential intermediary professional figure that dilutes the requirements to disclose some specific information about how AI systems work. This would create a situation in which, for instance, developers of fully automated wellness smartphone apps, but not physicians employing clinical-grade AI systems as aid to their activity, may have to deliver information about the use of AI systems to patients. On the other hand, in the case of fully automated systems, the communication requirements of the GDPR, may also add a further burden to doctors who are already overwhelmed by bureaucratic and administrative duties.

Furthermore, it should be reminded that, while our conceptual analysis arose from newly instituted GDPR provisions, it is not limited to medicals uses of AI, but applies to AI systems in general. As AI systems make inroads into multiple areas of human activity, including medicine, ensuring transparency requires addressing the opacity of such systems. With this article, we have proposed a conceptual framework to help shed light into the much-debated issue of AI opacity and the need for more transparent and accountable uses of artificial intelligence in medicine and beyond.

References

1. According to a widespread origin story, the term artificial intelligence (AI) was first introduced by mathematician and computer scientist John McCarthy in an application proposal for a Summer workshop held at Dartmouth College in 1956, which is considered the event that started AI as a scientific field (see Nils J Nilsson, The quest for artificial intelligence (Cambridge University Press 2009)). For more background on the research field and in-depth technical readings about AI we recommend the following two books:Stuart J Russell and Peter Norvig, Artificial intelligence: a modern approach (Pearson Education Limited 2016) and Ian Goodfellow, Yoshua Bengio and Aaron Courville, Deep Learning (MIT Press 2016).
2. Dzonale A Andrews, James Bonta and J Stephen Wormith, 'The Recent Past and Near Future of Risk and/or Need Assessment' (2006) 52 Crime & Delinquency 7.
3. Kori Hale, 'AI Challenging Bank Lending Practices' (Forbes, 2018) https://www.forbes.com/sites/korihale/2018/05/22/ai-challenging-banks-lending-practices/#77f499657e0e accessed 10 July 2018; Jon Walker, 'Artificial Intelligence Applications for Lending and Loan Management' (TechEmergence, 2018) https://www.techemergence.com/artificial-intelligence-applications-lending-loan-management/ accessed 10 July 2018.
4. Executive Office of the President and others, Big data: A report on algorithmic systems, opportunity, and civil rights (Executive Office of the President 2016); Ramnath Balasubramanian, Ari Libarikian and Doug McElhaney, 'Insurance 2030—The impact of AI on the future of insurance' (McKinsey, 2018) https://www.mckinsey.com/industries/financial-services/our-insights/insurance-2030-the-impact-of-ai-on-the-future-of-insurance accessed 10 July 2018.
5. Varun Gulshan and others, 'Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs' (2016) 316 Jama 2402.
6. B Ehteshami Bejnordi and others, 'Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer' (2017) 318 JAMA 2199.
7. Bernard Marr, 'First FDA approval for clinical cloud-based deep learning in healthcare' Forbes
8. FDA, 'FDA permits marketing of artificial intelligence algorithm for aiding providers in detecting wrist fractures' (FDA, 2018) https://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm608833.htm accessed 29 May 2018.
9. We use the term 'rule' in a general sense, and not as commonly used in the AI field, as for instance in the case of a 'rule-based expert system'. Therefore, throughout this paper, with the word 'rule' we refer to the operations the system undertakes in order to perform a task. In the case of AI systems employing ML algorithms, such 'rules' are autonomously learnt by the system – either with or without human supervision.
10. Mike Rawson, 'For artificial intelligence to thrive, it must explain itself' (The Economist, 15 February 2018) https://www.economist.com/science-and-technology/2018/02/15/for-artificial-intelligence-to-thrive-it-must-explain-itself; Niall Ferguson, 'The more machines learn, the less we shall grasp' (The Times, 11 February 2018) https://www.thetimes.co.uk/article/the-more-machines-learn-the-less-we-shall-grasp-x8kzddhxb accessed 13 July 2018; Will Knight, 'The Dark Secret at the Heart of AI' (MIT Technology Review, 11 April 2017) https://www.technologyreview.com/s/604087/the-dark-secret-at-the-heart-of-ai/ accessed 10 July 2018; Cliff Kuang, 'Can A.I. Be Taught to Explain Itself?' (The New York Times Magazine, 21 November 2017)

https://www.nytimes.com/2017/11/21/magazine/can-ai-be-taught-to-explain-itself.html accessed 10 July 2018.

11. Danton S Char, Nigam H Shah and David Magnus, 'Implementing machine learning in health care—addressing ethical challenges' (2018) 378 The New England journal of medicine 981.
12. ibid 983.
13. Sandra Wachter, Brent Mittelstadt and Luciano Floridi, 'Why a right to explanation of automated decision-making does not exist in the general data protection regulation' (2017) 7 International Data Privacy Law 76.
14. ibid 78.
15. ibid 78.
16. Andrew D Selbst and Julia Powles, 'Meaningful information and the right to explanation' (2017) 7 International Data Privacy Law 233, 239.
17. Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data.
18. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), OJ 2016 L 119/1 ('GDPR').
19. Bryan Goodman and Seth R Flaxman, 'European Union regulations on algorithmic decision-making and a "right to explanation"' (2017) 38(3) AI Magazine 50; Selbst and Powles (n 16).
20. Wachter, Mittelstadt and Floridi (n 13).
21. We therefore concur with Selbst and Powels that the GDPR endorses the principle that data subjects are entitled not only to be informed but also to receive some kind of meaningful explanation about the use of profiling and automated data-processing. This does not mean, however, that the scope of such demanded explanations is sufficient to dispel all types of opacity in the case of ML algorithms. In Section III, we argue that while the GDPR establishes the principle that some information about automated data processing and resulting decisions must be provided to data subjects, not all the sources of opacity are equally amenable to be removed. We also highlight that the rights and interests of data subjects (that in our case coincide with patients) are not equally affected or called into question by all forms of opacity. Therefore, while patients need transparency about automated data processing activities, they may not need a full disclosure of the inner workings of AI systems employed in the context of healthcare.
22. GDPR, recital 1.
23. ibid, recital 39.
24. ibid, recital 39.
25. ibid, recital 39.
26. The GDPR introduces or reaffirms a series of rights for data subjects, namely: the right to access (art 15), right to rectification (art 16), right to erasure (art 17), right to restriction of processing (art 18), right to data portability (art 20), right to object and automated individual decision-making (arts 21 and 22); ss 1 to 5 of ch IV specify the obligations and requirements for data controllers and processors (arts 24 to 43).
27. GDPR, art 4(4) (emphasis added).
28. However, it has to be stressed that other forms of data processing not employing AI could also amount to profiling according to the above definition.
29. Article 29 Data Protection Working Party, 'Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679' (2017) WP251rev.01 ('Guidelines').
30. ibid 7.

31. ibid 5.
32. GDPR, arts 5-s and 13-21.
33. ibid, arts 5-s and 13-21.
34. ibid, art 22.
35. In EU legislation, recitals, albeit not binding per se, provide indications about how to the subsequent articles have to be interpreted and what they aim to realize.
36. GDPR, recital 63.
37. ibid, recital 71.
38. ibid, art 22(1).
39. Other general principles applying to data processing of personal data, and therefore also to profiling an automated decision-making include purpose limitation (ie data can be used only for purposes coherent with the original purpose of collection, art 5(1)(b)), data minimisation (art 5(1)(c)), accuracy (art 5(1)(d)), storage limitation (art 5(1)(e)), data security (art 5(1)(f)), and accountability (art 5(2)).Informed consent by data subject can be a lawful condition for data processing activities, including profiling and non-solely automated decision making (art 6(1)(a)). The Guidelines, again, link this requirement to opacity (page 13). Since data processing can rely on data not directly provided by the data subject, inform consent, is a way to inform people about an envisaged activity regarding him- or herself and its foreseeable consequences (ibid). Consent should thus be considered a sufficient but not necessary condition for lawful data processing. Other equally sufficient conditions include: processing is necessary for the performance of a contract (art 6(1)(b)); processing is needed to comply to legal obligations (art 6(1)(c)); vital interests are at stake (art 6(1)(d)); processing is in the public interest or needed to exercise official authority (art 6(1)(e)); processing is necessary to fulfil a legitimate interest (art 6(1)(f)).
40. Whenever their data are collected, according to art 13(1) and (2), data subjects have a right to receive specific information about: data controllers; the purposes of processing; the location of processing; the legitimate interests of data controllers; the recipients of their data; the duration of processing; their rights; whether data processing is a statutory or contractual requirement, is needed to enter a contract, or is mandatory (and the possible consequences of not providing data). When data are not obtained directly by the data subject, information should also include: the categories of data concerned and the source of the personal data (art 14(1) and (2)). Additionally, data subjects have a right to access their collected data at any moment (art 15(1)). Data subjects also have a right to rectify their data (art 16), to have them erased (art 17), and to restrict the scope of processing (art 18). Furthermore, according to art 21(1), data subjects have a right to object to automated individual decision-making, including profiling in the case of direct marketing (art 21(2), art 21(3)), but also in the case processing by public authorities or by third parties unless the controller demonstrates compelling legitimate interests that override data subjects' rights and freedom (as specified in art 6(1)(e) and 6(1)(f)). The same art 21, on para 6 specifies that data subjects can also object to processing in the context of research unless there are public interest reasons to continue processing their data (art 21(6)).
41. GDPR, art 22(1). The Guidelines specify that a legal effect occurs if a decision affects a person's legal rights or status, or rights that emerge from a contract (Guidelines 21). According to the Guidelines 'similarly significant' effect is one that is of 'sufficiently great or important' (Guidelines 21). More concretely, they state that a 'decision must have the potential to: significantly affect the circumstances, behaviour or choices of the individuals concerned; have a prolonged or permanent impact on the data subject; or at its most extreme, lead to the exclusion or discrimination of individuals' (Guidelines 21). Processing of health data or of other personal data in the context of healthcare arguably produces an effect of this sort, since, at a minimum, it affects her circumstances and most likely her choices in a significant or otherwise important way.

42. ibid, art 22(4).
43. ibid, art 9(1).
44. ibid, art 13(2)(f).
45. ibid, art 14(2)(g).
46. ibid, art 15(1)(h); Guidelines 17.
47. ibid, art. 13(2)(f); ibid, art 14(2)(g); ibid, art 15(1)(h).
48. Frank Pasquale, The black box society: The secret algorithms that control money and information (Harvard University Press 2015).
49. Another extensive and similarly oriented analysis of the opacity concept is provided also by Burrel, who defines opacity either as intentional corporate or trade secrecy, as technical illiteracy, or as the way in which algorithms operate at the scale of application.
50. Article 29 Data Protection Working Party 5. The same articulation of opacity appears on page 9: 'profiling is often invisible to the data subject'.
51. ibid 13.
52. W Nicholson Price II, 'Black-box medicine' (2015) 28 Harv JL & Tech 419.
53. Diakopulous defines algorithmic opacity as both a lack of explanation and clarification of the AI system outcomes, as well as an extreme level of technical complexity that obfuscates the understanding of an AI system's working process (see Nicholas Diakopoulos, 'Algorithmic Accountability' (2015) 3(3) Digital Journalism 398). Finally, a group of researchers at Harvard differentiate between a type of opacity described as the unawareness of signals flowing through an AI system, and a type of opacity that consists in how certain factors were used by the AI to define a specific outcome (see Finale Doshi-Velez and others, 'Accountability of AI under the law: The role of explanation' arXiv preprint arXiv:171101134).
54. GDPR, recital 63.
55. ibid, arts 13, 14 and 15.
56. ibid, art 22(1).
57. ibid, arts 13(2)(f) and 15(1)(h).
58. ibid, art 14(2)(g).
59. ibid, arts 13 and 14.
60. Article 4(4) of the GDPR: '"profiling" means any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements;'.
61. See s IV.
62. Informed consent is anyway mandatory for solely automated employing special types of data, including health data, as specified in art 9 of the GDPR.
63. Effy Vayena and Alessandro Blasimme, 'Biomedical Big Data: New Models of Control Over Access, Use and Governance' (2017) 14 Journal of Bioethical Inquiry 501.
64. GDPR, arts 13, 14 and 15.
65. ibid, arts 13(2)(f), 14(2)(g) and 15(1)(h).
66. Article 29 Data Protection Working Party 17.
67. Gulshan and others (n 5).
68. Kun-Hsing Yu and others, 'Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features' (2016) 7 Nature communications 12474.
69. Ziad Obermeyer and Ezekiel J Emanuel, 'Predicting the future—big data, machine learning, and clinical medicine' 375 The New England journal of medicine 1216.
70. W Nicholson Price II, 'Medical Malpractice and Black-Box Medicine' Health Law, and Bioethics (Cambridge University Press, Forthcoming).
71. ibid.

72. Jenna Burrell, ‘How the machine "thinks": Understanding opacity in machine learning algorithms’ (2016) 3 Big Data & Society 1.
73. Matthew D Zeiler and Rob Fergus, Visualizing and understanding convolutional networks (Springer 2014).
74. As Hildebrandt (see Mireille Hildebrandt, ‘Defining profiling: a new type of knowledge?’ in Profiling the European citizen (Springer 2008)) has explained, ‘correlations stand for a probability that things will turn out the same in the future. What they do not reveal is why this should be the case’ (see Hildebrandt, page 18). The use of algorithmic decisions in an increasingly wider range of applications has led some (see Hildebrandt, eg page 27) to caution against the rise of a ‘black box’ society and demand increased transparency in algorithmic decision-making.
75. Peter Machamer, Lindley Darden and Carl F Craver, ‘Thinking about Mechanisms’ (2000) 67 Philosophy of Science 1.
76. Ana Ferreira and others, ‘Sickle Hemoglobin Confers Tolerance to Plasmodium Infection’ (2011) 145 Cell 398.
77. Burkhard Hinz, Olga Cheremina and Kay Brune, ‘Acetaminophen (paracetamol) is a selective cyclooxygenase-2 inhibitor in man’ (2008) 22 The FASEB journal 383.
78. Tanya Lewis, ‘Mystery Mechanisms’ (The Scientist, 29 July 2016) https://www.the-scientist.com/news-analysis/mystery-mechanisms-33119 accessed 1 July 2018.
79. Article 29 Data Protection Working Party 25-26.
80. Sandra Wachter, Brent Mittelstadt and Chris Russell, ‘Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR’ (2017) 31 Harvard Journal of Law & Technology 6.
81. Doshi-Velez and others (n 53)
82. Effy Vayena and Alessandro Blasimme, ‘Health Research with Big Data: Time for Systemic Oversight’ (2018) 46 The Journal of Law, Medicine & Ethics 119.
83. Ryen W White, P Murali Doraiswamy and Eric Horvitz, ‘Detecting neurodegenerative disorders from web search signals’ (2018) 1 npj Digital Medicine 8.
84. Laura Weiss Roberts, Steven Chan and John Torous, ‘New tests, new tools: mobile and connected technologies in advancing psychiatric diagnosis’ (2018) 1 npj Digital Medicine 20176.
85. Gillian Gresham and others, ‘Wearable activity monitors to assess performance status and predict clinical outcomes in advanced cancer patients’ (2018) 1 npj Digital Medicine 27.
86. Effy Vayena and others, ‘Policy implications of big data in the health sector’ 96 Bulletin of the World Health Organization 66.
87. Alessandro Blasimme and Effy Vayena, ‘Becoming partners, retaining autonomy: ethical considerations on the development of precision medicine’ (2016) 17 BMC Medical Ethics 67; Alessandro Blasimme and Effy Vayena, ‘"Tailored-to-You": Public Engagement and the Political Legitimation of Precision Medicine’ (2016) 59 Perspectives in Biology and Medicine 172.
88. Effy Vayena and others, ‘Digital health: meeting the ethical and policy challenges’ (2018) 148 Swiss Medical Weekly.
89. Alessandro Blasimme and others, ‘Data Sharing For Precision Medicine: Policy Lessons And Future Directions’ (2018) 37 Health Affairs 702.
90. M Simpson and others, ‘Doctor-patient communication: the Toronto consensus statement’ (1991) 303 British Medical Journal 1385.
91. Glyn Elwyn and others, ‘Shared Decision Making: A Model for Clinical Practice’ (2012) 27 Journal of General Internal Medicine 1361.
92. Wachter, Mittelstadt and Floridi (n 13).

CHAPTER 4: Mobile Apps for Travel Medicine and Ethical Considerations: a Systematic Review

Submitted for review as: Ferretti, A.*, Hedrich, N.*, Lovey, T.*, Vayena, E., Schlagenhauf, P. Mobile Apps for Travel Medicine and Ethical Considerations: a Systematic Review.

* These authors contributed equally and share first authorship

4.1 Abstract

The advent of mobile applications for health and medicine will revolutionize travel medicine. Despite their many benefits, such as access to real-time data, mobile apps for travel medicine also come with ethical issues, including questions about security and privacy. This systematic literature review examined the ethical considerations of mobile apps for travel medicine. Database screening yielded 1795 results and 7 papers satisfied the criteria for inclusion. A mixture of inductive and deductive data extraction examined both the benefits and challenges, as well as ethical considerations, of mobile apps for travel medicine. Ethical considerations were discussed with varying depth across the included articles, with *privacy* and data protection mentioned most frequently, highlighting concerns over sensitive information and a lack of guidelines in the digital sphere. Additionally, this review highlights the scarcity of discussion around ethical issues, and the need for greater consideration of ethics in each step of app development and use.

4.2 Introduction

Travel, whether for leisure, business, or visiting friends and relatives (VFR) is an important global phenomenon, with significant impacts on spending, employment, and also health. In 2019, there were 1.5 billion international inbound tourists, with Europe having the largest number of international tourists and the most spending on tourism (1, 2). With the growth of international travel, however, comes an increased risk to traveler health, and of the possibility of the spread of infections to new areas. Travelers may be at risk of contracting illnesses such as malaria, traveler's diarrhea, arboviruses (such as dengue, Zika, and chikungunya), sexually transmitted infections, and more recently, the novel coronavirus 2019 SARS-CoV-2 (3-6).

Travel medicine plays an important role in preventing and treating travel-related illnesses. In Europe, travel medicine is a diverse field with a variety of national and local guidelines, and is administered by a wide range of health professionals, including nurses, general practitioners, travel clinics, and pharmacists (2). Prevention is key for maintaining traveler health, and can include vaccinations, prophylaxis, travel safety information, insect bite prevention, and more (7). Also relevant is the role of travelers as sentinels for infection and in surveillance of imported infections associated with travel. As travel increases and diversifies in destinations, and numbers and types of travelers, so too must travel medicine respond to the changing landscape of travel.

One method that has shown promise is the use of smartphone apps, or mHealth apps (8). Monitoring traveler health behavior as well as encountered risks has become easier and more reliable due to advances in the quality of mobile health technology and widespread use of smartphones, allowing for real-time data collection (9, 10). An ambitious new project called Illness Tracking in Travellers (ITIT) aims to collect data on traveler illness in collaboration with the World Health Organisation (WHO), with a goal of facilitating rapid public health responses(11).

However, many travel medicine apps are not up to date, lack accurate and evidence-based content, or were not developed with the involvement of health professionals (12). This is consistent with the broader literature on health apps (13-15). Research has shown that questions of data security, confidentiality, liability, and trust are at the forefront of the discussion about health apps, despite their many advantages (13, 16, 17). Effectiveness and accessibility are also mentioned frequently as reasons for the use or rejection of health apps (18, 19). Equity of access is another important ethical issue. Although the average number of mobile phone subscriptions worldwide was 104 per 100 people in 2018 (20), certain populations are underrepresented, including older individuals and those with a lower socioeconomic status (21, 22). This information is particularly relevant for studies of travel health apps: despite their intention to collect information from a variety of settings and population groups, these studies might be biased towards subgroups already owning and comfortably using mobile devices (23). These issues are important to address in order to avoid bias. User trust is another important issue and lack of trust can result in poor uptake(24).

The goal of this systematic review is to evaluate ethical issues around mobile health apps for travelers, identify important deficits, and suggest key ethical areas to address in future travel medicine apps.

4.3 Methods

Identification and Selection of Studies

The systematic review was conducted in accordance with PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) guidelines (25) and registered in the Prospero database (CRD42021231857). A systematic search of the Medline, Scopus, Web of Science, Embase, IEEE Xplore, Science Direct, Cochrane Central Register of Controlled Trials (CENTRAL), SSRN, and medrXiv databases was performed on January 7th, 2021 by a librarian

scientist. The search strings can be seen in Appendix 1 Chapter 4. Titles and abstracts were imported into the reference manager software Endnote20 (Clavirate, 36T3 Boston, MA 02210), and duplicates were removed. Titles and abstracts were then imported into the knowledge synthesis software Rayyan QCRI (26) and examined for eligibility by two independent reviewers, with the consultation of a third in case of disagreement. Finally, the full text of the remaining studies was examined for relevance, and relevant studies were included in this review. The reference lists of included papers were examined for additional relevant studies not included in the initial search. A team of three co-authors completed the abstract screening, full-text review, and data extraction. Any disagreement among the authors was resolved through discussion.

Eligibility Criteria

Only studies meeting the inclusion criteria were considered. Reviewed studies were written in English, German, French, or Italian, and published until the 31st December 2020. Preprints, dissertations, and peer-reviewed studies with all study designs (qualitative, mixed methods, quantitative) were included, while books, conference abstracts, editorials, and papers without an available full text were excluded. Duplicates and irrelevant papers were also excluded. In order to be considered relevant, papers had to report on mobile phone apps for travel medicine for travelers over 18 (international and intranational), and these apps must have been developed for the primary purpose of traveler health/travel medicine. Apps for children and youth were excluded, as well as apps not designed specifically for travel medicine, even though they may still collect data useful for travel medicine research (such as social media apps collecting epidemiological data), or may be used in some way by travelers (such as holiday booking apps, apps for tourist leisure activities). Reference to ethical implications of developing and using mobile applications for travel medicine was an additional inclusion criterion. Reasons for exclusion from the review were noted in Rayyan QCRI (26).

Data Extraction

The primary outcome was ethical considerations of the development and use of mobile phone apps for travel medicine purposes, and the secondary outcome was the opportunities and challenges in ethical considerations. Relevant information was extracted through a deductive coding process. In consultation, all authors agreed on a list of categories to code the studies accordingly. When an ethical consideration included in the text could not be coded under any existing category, it was temporarily designated "unclassified". Subsequently, the authors decided whether this code should generate a separate category (introduced through an inductive process)

or be grouped under an existing one. The extracted information was presented in tabular form using Excel software.

Risk of Bias Assessments

Quality assessment of the studies was conducted simultaneously. At the study level, quality was assessed with different tools according to the study design (Randomized trials – Cochrane risk of bias tool, Observational studies – STROBE, Narrative articles – SANRA)(27). At the outcome level, we assessed the types of reasons supporting each ethical statement: supported by empirical evidence, justified by rationally articulated arguments (potentially supported by the literature), or uncorroborated (without an explicit justification). This categorization allowed for higher precision in identifying the gaps in the ethical reflection on travel medicine apps (28). The quality assessment (recorded in Appendix 2 Chapter 4) was once again conducted independently by two authors, and disagreement was resolved through discussion with a third.

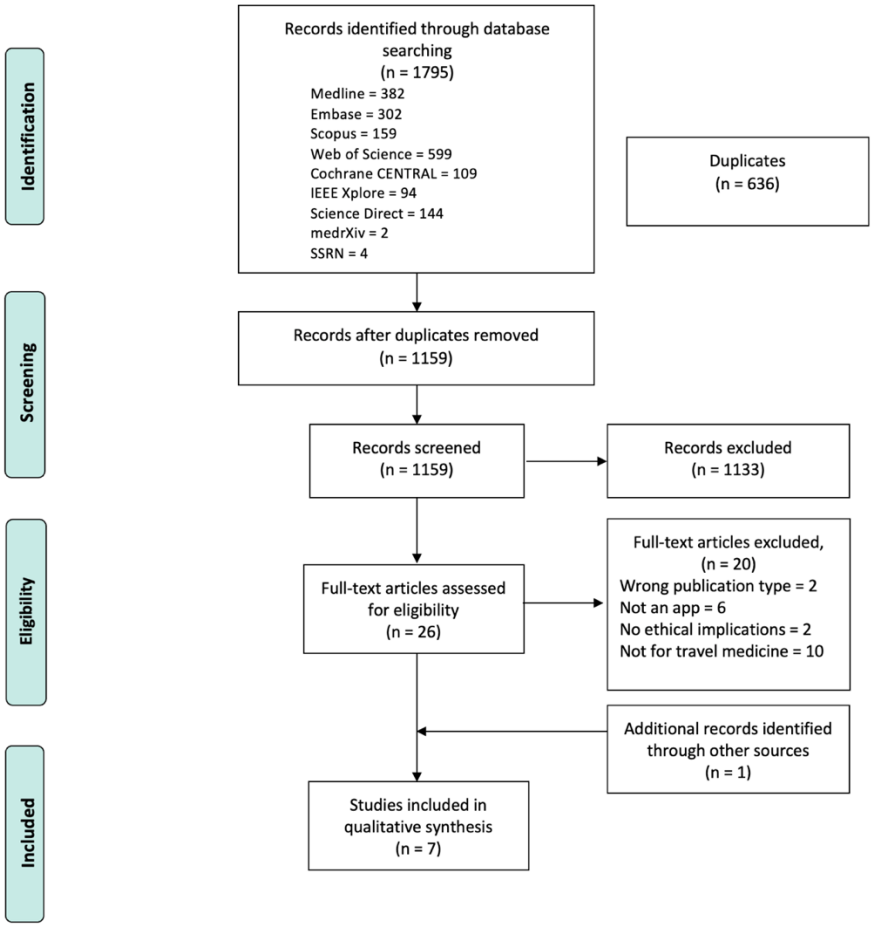
Data Synthesis

All papers that met the eligibility criteria were included in the narrative synthesis (29). Similarities and differences across studies were analyzed, and homogeneous studies were clustered. Study characteristics, type of intervention adopted, context of the intervention, opportunities and challenges brought by the intervention, and ethical considerations of developing and adopting mobile apps for travel medicine purposes were all considered in the synthesis. As a qualitative synthesis, the findings were clustered thematically according to the reasons used to justify the ethical considerations.

4.4 Results

A total of 1795 studies were found through the literature search. Of these, 636 were duplicates, and 1133 were excluded through the abstract screening. The full text of the remaining 26 papers were screened, and of these, 6 were included. In addition, 1 paper was found through the screening of reference lists of the included papers, resulting in 7 papers being included in the review. Figure 4.1 provides an overview of the screening process.

Figure 4.1: PRISMA flowchart of identification and selection of studies to be included in the systematic review



Of the 7 included papers, 2 were cohort studies and 5 were qualitative analyses or narrative reviews. Characteristics of included papers can be found in Table 4.1. The two cohort studies described the same app called the *Tourist* app, which was pilot tested in the 2018 paper. (30) The 2020 paper focuses on novelties and upgrades of the app, as well as participant willingness to use the app. Three papers described specific apps for travel medicine: Du et al. (contact tracing) (31), Subramaniyaswamy et al. (food recommendations while travelling) (32), and Sethia et al. (electronic health record access while travelling) (33). Finally, two papers provided a review of several apps. Seed et al. (12) offered an overview of travel medicine apps available in 2016, and Lai et al. (34) reviewed the literature on benefits and challenges of travel medicine mHealth.

Table 4.1: Characteristics of Included Papers

Author	Year	Title	Journal	Study Type	Field
Baroutsou et al.	2020	TOURIST2 – Tracking of urgent risks in swiss travelers to the 6 main travel destinations – Feasibility and ethical considerations of a smartphone application-based study	Travel Medicine and Infectious Disease	Cohort Study	Epidemiology
Farnham et al.	2018	Streaming data from a smartphone application: A new approach to mapping health during travel	Travel Medicine and Infectious Disease	Cohort Study	Epidemiology
Du et al.	2020	COVID-19 Contact Tracing Apps: A Technologic Tower of Babel and the Gap for International Pandemic Control	JMIR MHealth and UHealth	Qualitative Analysis	Epidemiology
Lai et al.	2019	Measuring mobility, disease connectivity and individual risk: a review of using mobile phone data and mHealth for travel medicine	Journal of Travel Medicine	Qualitative Analysis	Epidemiology
Subramaniyaswamy et al.	2018	An ontology-driven personalized food recommendation in IoT-based healthcare system	Journal of Supercomputing	Qualitative Analysis	Computing
Sethia et al.	2018	Smart health record management with secure NFC-enabled mobile devices	Smart Health	Qualitative Analysis	Travel Medicine
Seed et al.	2016	Identification and review of mobile applications for travel medicine practitioners and patients	Journal of Travel Medicine	Brief Communication/ Qualitative Analysis	Travel Medicine

All included papers were rated for quality using the STROBE guidelines for the cohort studies, and the SANRA guidelines for the qualitative/narrative analyses (Table 4.2). The two cohort studies and the paper by Lai et al. (9) had the highest quality ratings, while the papers by Seed et al. (12) and Subramaniyaswamy et al. (32) had lower scores.

Table 4.2: Quality rating of papers included in the systematic review

Paper	Quality Rating System	Quality Rating
Baroutsou et al. 2020	STROBE[a]	21/22
Farnham et al. 2018	STROBE[a]	20/22
Du et al. 2020	SANRA[b]	10/12
Lai et al. 2019	SANRA[b]	11/12
Subramaniyaswamy et al 2018.	SANRA[b]	9/12
Sethia et al. 2018	SANRA[b]	10/12
Seed et al. 2016	SANRA[b]	9/12

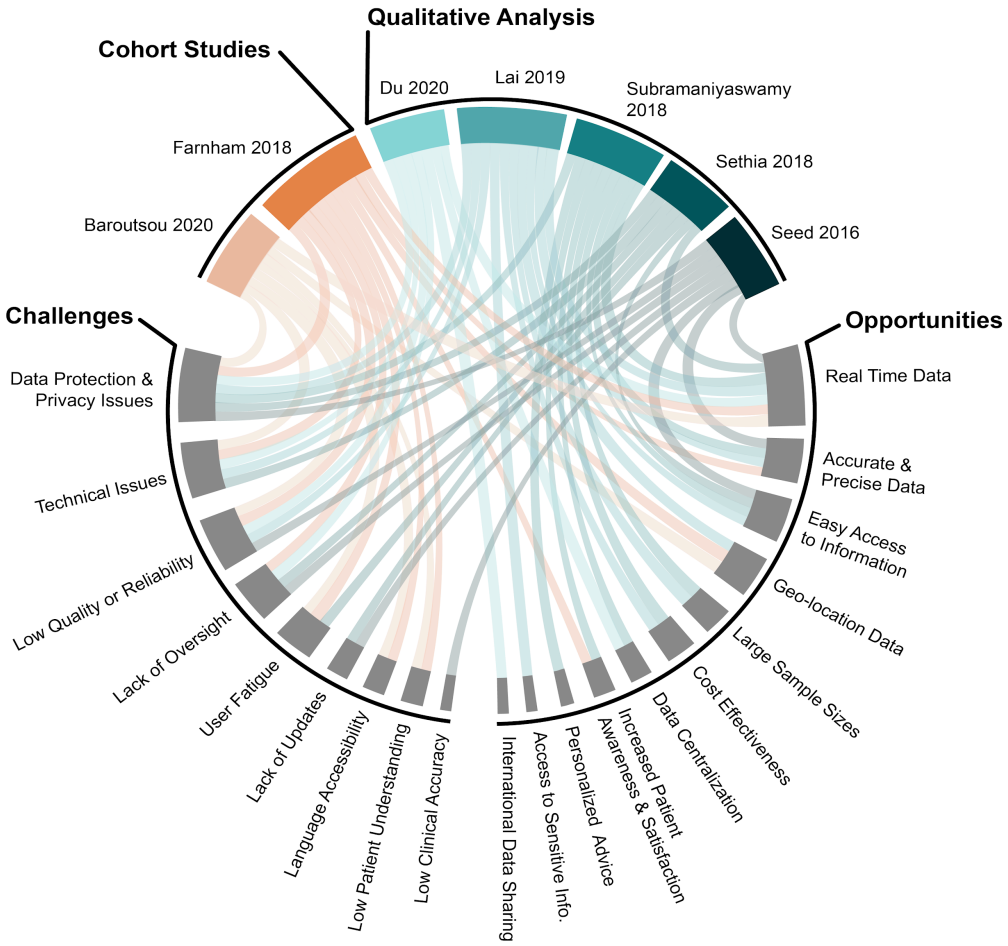
[a] Strengthening the Reporting of Observational studies in Epidemiology (STROBE) checklist was created to help authors submit high-quality observational studies by grading them on a total scale of 22 points.
[b] Scale for the Assessment of Narrative Review Articles (SANRA) aimed to improve the quality of narrative reviews by rating them on a 12-point scale.

Benefits and Challenges

Each paper mentioned opportunities and challenges of using mobile apps for traveler health, with reference to travel app users, researchers, and developers (Figure 4.2). The most commonly stated opportunity of travel medicine apps was to collect real-time data, thereby reducing recall bias and allowing users to access resources when needed. This was followed closely by the accuracy and precision of the data and easy access to information and resources, which are also related to reduced recall bias. Several papers mentioned linked geolocation data as a benefit of the apps, as well as the possibility of larger sample sizes and reduced costs. Geolocation benefits both researchers, enabling them to link location to risk events (34) or examine contact between users (as in COVID tracking apps) (31), and users, allowing for personalized information based on location (32). Finally, opportunities mentioned once or twice included personalized advice,

data decentralization, and easier international data sharing. Conversely, all of the papers recognized data protection and privacy issues as a challenge for travel health apps. Other potential weaknesses included technical issues, low-quality data, and low reliability. The lack of clear governance or oversight during app development was also highlighted as troublesome. Frequently mentioned challenges associated with mobile travel health apps included potential for user fatigue due to data overload, language accessibility concerns, lack of updates leading to outdated information, and low traveler understanding. The mentioned opportunities and challenges of mobile apps for traveler health are presented in Figure 4.2.

Figure 4.2: Challenges and opportunities identified for mobile apps used for travelers' health



General Ethical Issues

In five of the seven papers, a full section was dedicated to discussion of ethical issues, while two papers discussed ethical issues only briefly, devoting less than a paragraph to the topic. Sixteen distinct ethical issues were touched upon across all papers. However, despite the emphasis on ethical considerations, almost half were not explored in detail, with no justification of their relevance provided in the text. Instead, many issues were mentioned in passing in the methods section (Figure 4.3). More recently published papers tended to discuss a greater number of ethical issues and examine them in more detail than those published a few years ago. Furthermore, the recently published papers were more likely to contain evidence-based justification or stronger theoretical arguments in support of their ethical reasoning, in comparison with the older papers (Figure 4.4). In fact, only the cohort study by Baroutsou et al. (8) had evidence-based reasons

concerning topics such as secondary use of data, institutional trust, and age or chronic disease status of participants.

Looking more closely at the ethical considerations mentioned, privacy issues were most frequently discussed, being addressed by all of the papers, followed closely by issues included in the “CANDALS” classification (35) (Citizenship, Ability, Neurotypicality/Neurodiversity, Disability, Age, Literacy and/or fluency, and Size, BMI, or body habitus.) The papers discussed how age, disease status, ethnicity, lower-income country status, and health literacy can impact the adoption and usability of mobile health apps by individuals across countries, social classes, and cultures. Another frequently mentioned ethical issue was data storage, in relation to both data security (risk of cyber-attacks) and efficiency (e.g., saving energy in resource limited settings). Conversely, the least discussed ethical issues included transparency, autonomy, and individual traveler empowerment.

Figure 4.3: Cleveland's Dot Plots of 16 ethical considerations identified in the papers included in the systematic review

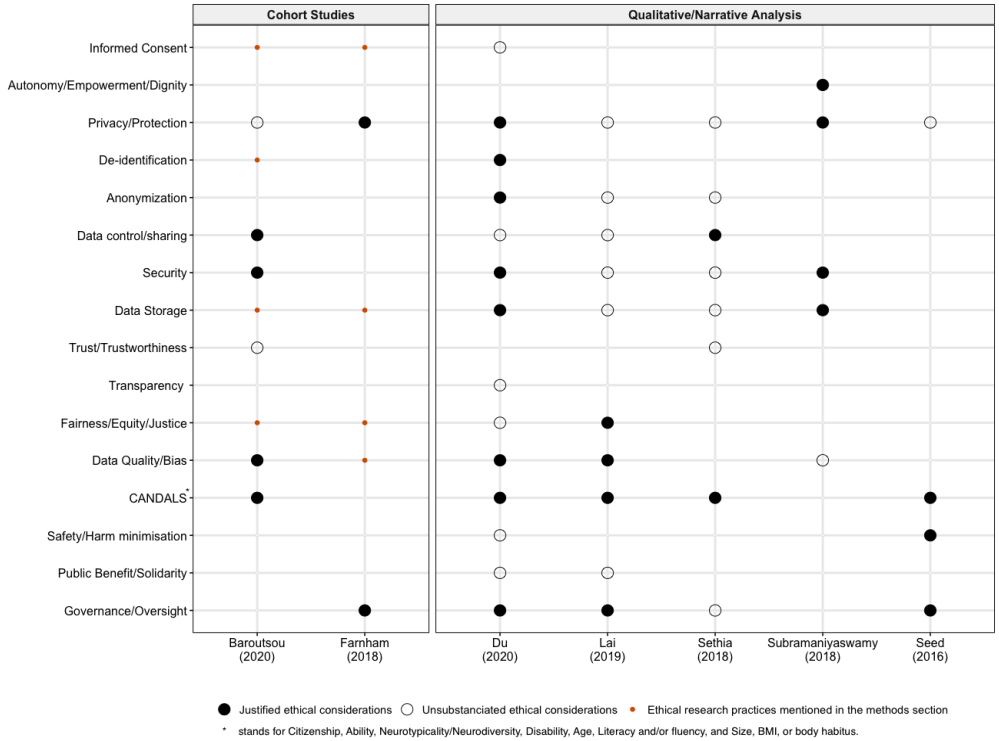
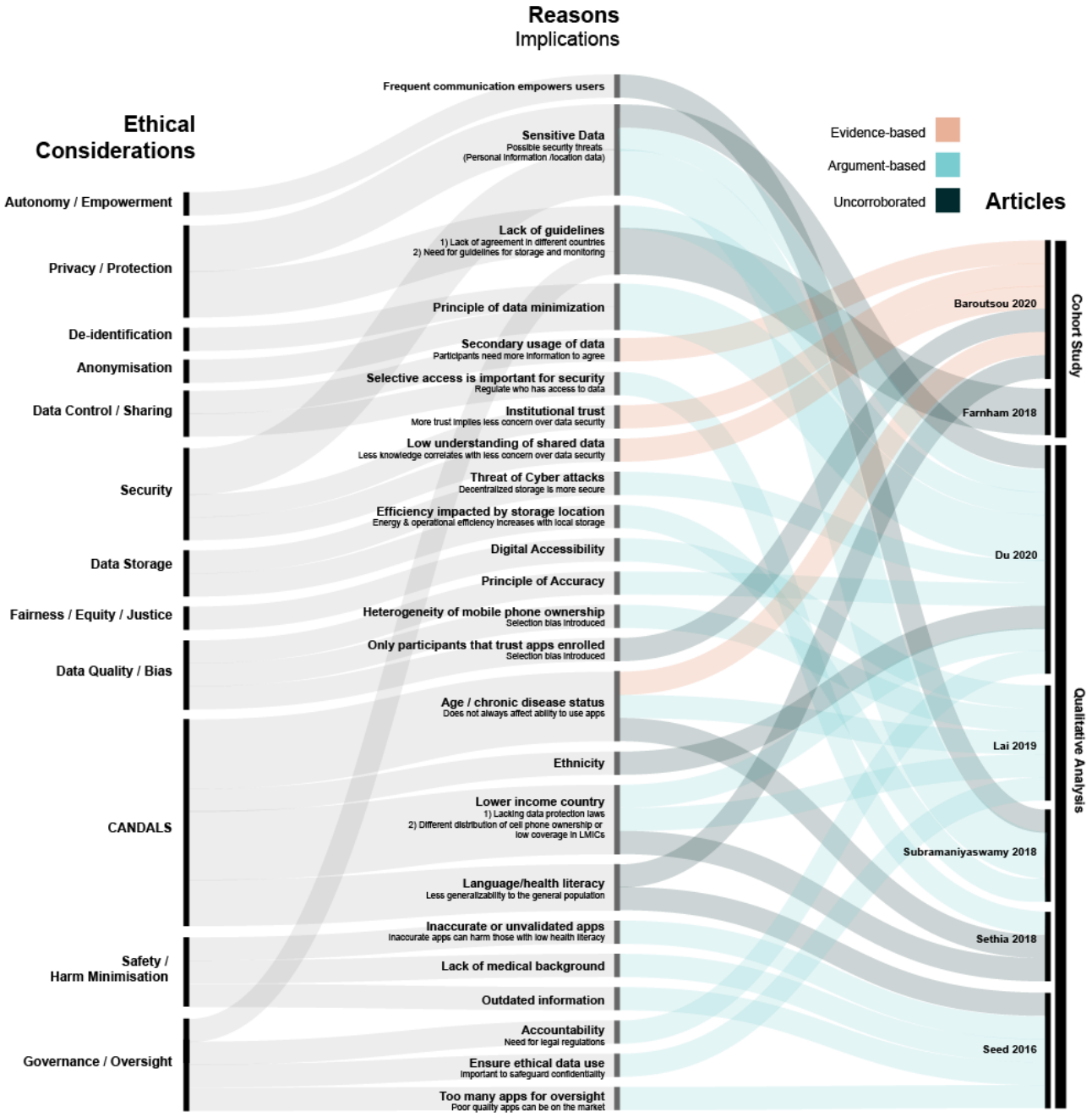


Figure 4.4: Types of reasons justifying the ethical considerations and their implications



Cohort Studies

Both cohort studies focused on the importance of protecting user privacy, an increasingly relevant topic for the general public. The 2018 paper by Farnham et al. (34) highlighted the lack of clear guidance at an international level, rendering it difficult to develop apps compliant with privacy laws across countries. The 2020 paper by Baroutsou et al. (8) goes beyond privacy issues, discussing the ethical implications of sharing data for secondary purposes, through surveying participant opinions of this topic before and after the study, and examining their reasoning. This highlights the importance of trust in the institutions responsible for app development, to engage app users and address data security concerns. Looking at additional ethical issues taken into account in the research methodology, Baroutsou et al. (8) mentioned e-consent forms and data de-identification and storage, as well as the concept of fairness, e.g., providing mobile devices to participants without access to one. They reflected further on data bias, as only individuals already interested in the app took part in the cohort study.

Qualitative/Narrative Papers

The paper by Du et al. (31) highlighted the greatest number of ethical concerns of all the included papers. Particularly, the paper discussed the data de-identification and anonymization as ways to preserve user privacy, as per the principle of data minimization. This paper also mentioned data security with regard to collection, storage, and use of sensitive data, as well as to individual harms that could emerge from a data breach. More specifically, it examined GPS location data used by apps such as those developed for COVID tracing, and the harms related to the potential theft of this information. The qualitative papers mention ethical considerations not considered in the cohort studies, such as transparency, public benefit, solidarity, safety, and harm minimization. Concerning the last point, Seed et al. mentioned inaccuracy, lack of a medical background, and outdated information due to a lack of updates as potential sources of harm for people using travel medicine apps, especially those with low health literacy. Sethia et al. (33) examined data control, emphasizing the importance of selective access to data for data security, and the importance of regulating data access. Furthermore, data quality (and its link to bias) was a concern mentioned by Du, Lai, and Subramaniyaswamy (9, 31, 32). Inaccurate data collection or heterogeneity of mobile phone ownership may result in selection bias, which can negatively affect data analysis and provide inaccurate feedback to users. The majority of the qualitative papers also mentioned issues of data governance, specifically the lack of adequate oversight for mHealth apps in the field of travel medicine. Du et al. stressed the need for legal regulation to address accountability, (31) in the case of a security breach or inaccurate recommendations made by an app. Lai et al.

(9) recommended introducing oversight to ensure that privacy is taken into account during travel medicine app development. Finally, Seed et al. (12) reflected on the exponential number of apps developed in recent years, and the (lack of) effectiveness of current oversight mechanisms to keep pace with this rapidly evolving sector.

4.5 Discussion

This review found that privacy is the most pressing ethical issue for travel medicine apps. This may be partially explained by researcher and developer concerns about compliance with privacy and security regulations. These concerns are justified, due to the lack of clear ethical standards and data regulation at the international level (36). Apart from the General Data Protection Regulation in Europe, there are no defined minimum global standards for storage and sharing of personal data for secondary purposes (37, 38). Medical travel apps (as all health apps) must comply with each individual country's privacy law (34). Baroutsou et al. (8) showed that trust in the institutions developing and implementing health apps can reduce user fears about data security and confidentiality. It is therefore essential to develop international data governance standards, endorsed by a variety of stakeholders, that not only guide researchers when developing their applications, but also increase user trust in the technology (39).

Given the types of papers assessed (cohorts and papers describing app development) it is not surprising that data quality and bias were also predominant issues. As the papers were written from the perspectives of app developers and researchers, concerns about potential biases and other technical issues were highlighted over issues that might have been emphasized by ethicists. Examining data quality in more detail, self-reported user data introduces two issues of ethical relevance. The first is data accuracy. Although real-time self-reporting of data can reduce recall bias, positively influencing data quality, researchers can struggle to verify whether the information provided is precise, complete, and mirrors reality. For this reason, using GPS and metadata collected directly through the phone (without user input) might compensate for potential errors and biases. The ability to access these data represents a significant advantage of travel applications over other travel medicine strategies. Nevertheless, rigorous data quality control is still required. The second issue is data representativeness. Our analysis showed that effort should be made to include minorities as well as other population subgroups (CANDALS) in the design and deployment of health apps, as factors such as age, language and health literacy, or living in a lower-middle income country play a role in app use (9, 33). Selection bias introduced due to the

heterogeneity of mobile phone ownership or user comfort with mobile technology directly affects data quality. This in in turn may give incorrect or misleading feedback to users, which is particularly problematic for travel medicine apps, when user health is at stake.

Conversely, researchers dedicated only minimal attention to issues of equity and justice. Although a few articles (9, 34) discussed accessibility through lending a mobile phone or SIM card to participants, no reference is made to the social implications of these applications, or whether they extend access to health information in an equitable way to all population groups. Similarly, though the apps are used by individuals with various needs and health concerns, it can seem that researchers developed these tools without adequately considering the context, resulting in a "one size fits all" application. Only the more recent cohort study considered engaging users in the app development process and receiving feedback. Following on this point, it is important to note that informed consent, central to mHealth literature, has scarcely been discussed. Informed consent is mentioned in the cohort study methodology without further development, though their protocols reference it often. Of the qualitative studies, informed consent is only briefly referenced in Du et al. (31). Many of the papers seem to view informed consent more as a task to be completed to avoid legal repercussions, than as a real ethical concern. However, in the interest of increasing trust, researchers should engage users. This might include clearly communicating the app's objectives and addressing the data confidentiality concerns of users. Moreover, researchers should focus on user satisfaction, providing an app that is intuitive and accessible on multiple platforms. Finally, it could be important not only to state the user's benefits from the app but also to stress the benefit for the broader community (as with COVID-19 digital contact tracing apps). If researchers succeed in increasing willingness to use the app, they may also indirectly increase the quantity and quality of data that they collect.

Accompanying the user on their journey, travel medicine apps can offer individualized advice although this would mean that the app becomes a "medical device" and would thus require regulation. However, whether or not these apps are actually effective in providing timely advice and suggestions was not discussed in the papers evaluated here. On the contrary, as pointed out in Seed et al. (12), there is potential for harm due to a lack of medical background of app developers and app users and poor data accuracy. This should be considered carefully by researchers, as it may negatively influence user willingness to adopt the apps, especially those that collect highly sensitive data (16). More research is needed to evaluate the ethical and societal implications of travel medicine apps.

4.6 Strengths and Limitations

This is the first systematic review to examine the important and quickly growing topic of ethical aspects of travel medicine apps. A major strength of this work is the evaluation of key health equity stratifiers using the CANDELS classification to show how age, disease status, ethnicity, lower-income country status, and health literacy can impact the adoption and usability of mobile health apps by individuals across countries, social classes, and cultures. In the modern age, digital technology will play an expanding role in travel, emphasizing the importance of analyses such as this one. One limitation of this review is the quality of ethical assessment within the selected papers. Although the 7 included papers matched the inclusion criteria and were of good quality, the depth of ethical assessment was often superficial, with only a short section devoted to ethics and little evidence to support the issues discussed. This reinforces the need for more research into ethical issues surrounding travel medicine apps, and health apps in general. Another limitation of this analysis is the inclusion and analysis of both cohort and qualitive studies, even though they employ different methodologies. However, the discussion of ethical issues can occur across all study designs, leading to comparability, and the quality assessment of the selected studies indicates strong results across study types. A final limitation is the use of a qualitative thematic methodology to extract ethical issues. This procedure might be subject to subjective biases, which were addressed by 1) having an inductive table of ethical issues and using a deductive approach to collect the issues, and 2) having multiple researchers working in parallel. However, it is not possible to completely rule out bias in the data extraction.

4.7 Conclusion

This systematic review identified 1159 unique articles of which 7 (0.6%) met our pre-defined inclusion criteria. We found that although some ethical issues are widely debated (privacy, security and data quality), many are just mentioned (justice, fairness, risk assessment), and some are disregarded (effectiveness, user involvement). While it is true that travel applications constitute a relatively new approach to collecting data and engaging users, this result revealed gaps that exist regarding ethical considerations in travel medicine literature. These gaps highlight the need for developers and researchers working with travel medicine apps to do a careful risk-benefit assessment, not only exploring potential risks, but employing strategies to mitigate such risks. In light of the fast-evolving landscape of digital health and health apps, oversight

mechanisms should be updated to support researchers and developers in making ethically aligned choices.

Acknowledgments

We thank Dr. Sabine Klein for her valuable help and support in defining the search string and performing the literature search on the relevant databases. We are grateful to Joanna Sleigh for her support in designing Figures 2 and 4, and to Shannon Hubbs for editorial suggestions.
This work was supported by the Swiss National Science Foundation funded research project, “Illness Tracking in Travellers (ITIT)”. Grant number: 320030_192653.

Author Contributions

AF, NH and TL selected the articles, compiled the results, drafted the manuscript. All authors participated in the design of the project, the development of the final manuscript, and approved the submitted version.

Competing Interests

None declared

References

1. International tourism growth continues to outpace the global economy | UNWTO 2021 [Available from: https://www.unwto.org/international-tourism-growth-continues-to-outpace-the-economy.
2. Schlagenhauf P, Santos-O'Connor F, Parola P. The practice of travel medicine in Europe. Clin Microbiol Infect. 2010;16(3):203-8.
3. Kuenzli E. Antibiotic resistance and international travel: Causes and consequences. Travel Med Infect Dis. 2016;14(6):595-8.
4. Tatem AJ, Jia P, Ordanovich D, Falkner M, Huang Z, Howes R, et al. The geography of imported malaria to non-endemic countries: a meta-analysis of nationally reported statistics. The Lancet Infectious Diseases. 2017;17(1):98-107.
5. Memish ZA, Osoba AO. Sexually transmitted diseases and travel. International Journal of Antimicrobial Agents. 2003;21(2):131-4.
6. Stoddard ST, Morrison AC, Vazquez-Prokopec GM, Paz Soldan V, Kochel TJ, Kitron U, et al. The role of human movement in the transmission of vector-borne pathogens. PLoS Negl Trop Dis. 2009;3(7):e481.
7. Korzeniewski K. Travel health prevention. Int Marit Health. 2017;68(4):238-44.
8. Baroutsou V, Hatz C, Blanke U, Haile SR, Fehr J, Neumayr A, et al. TOURIST2 – Tracking of urgent risks in swiss travellers to the 6 main travel destinations – Feasibility and ethical considerations of a smartphone application-based study. Travel Medicine and Infectious Disease. 2021;39.
9. Lai S, Farnham A, Ruktanonchai NW, Tatem AJ. Measuring mobility, disease connectivity and individual risk: a review of using mobile phone data and mHealth for travel medicine. J Travel Med. 2019;26(3).
10. Chiodini J. TravWell--The latest CDC app for travellers. Travel Med Infect Dis. 2015;13(3):276-7.
11. Ponce C, Dolea C. The World Health Organisation (WHO) and International Travel and Health: New collaborative, evidence-based and digital directions. Travel Med Infect Dis. 2019;27:1.
12. Seed SM, Khov SL, Binguad FS, Abraham GM, Aungst TD. Identification and review of mobile applications for travel medicine practitioners and patients. J Travel Med. 2016;23(4).
13. Martinez-Martin N, Kreitmair K. Ethical Issues for Direct-to-Consumer Digital Psychotherapy Apps: Addressing Accountability, Data Protection, and Consent. JMIR Ment Health. 2018;5(2):e32.
14. Powell AC, Landman AB, Bates DW. In search of a few good apps. JAMA. 2014;311(18):1851-2.
15. Larsen ME, Huckvale K, Nicholas J, Torous J, Birrell L, Li E, et al. Using science to sell apps: Evaluation of mental health app store quality claims. NPJ Digit Med. 2019;2:18.
16. Flaherty GT. Research on the move: the potential applications of mobile health technology in travel medicine. Journal of Travel Medicine. 2016;23(6):taw061.
17. Yang YT, Silverman RD. Mobile health applications: the patchwork of legal and liability issues suggests strategies to improve oversight. Health Aff (Millwood). 2014;33(2):222-7.
18. Stephan LS, Dytz Almeida E, Guimaraes RB, Ley AG, Mathias RG, Assis MV, et al. Processes and Recommendations for Creating mHealth Apps for Low-Income Populations. JMIR Mhealth Uhealth. 2017;5(4):e41.
19. Ahmed OH, Lee H, Marchant H, Jones R, Hall EE. Navigating the new landscape of apps: Overcoming the challenge of poor quality apps in sport and exercise medicine. Br J Sports Med. 2018;52(20):1283-4.

20. Mobile cellular subscriptions (per 100 people) | Data 2021 [Available from: https://data.worldbank.org/indicator/IT.CEL.SETS.P2?end=2019&name_desc=false&start=1980&type=shaded&view=chart.
21. Wesolowski A, Buckee CO, Engo-Monsen K, Metcalf CJE. Connecting Mobility to Infectious Diseases: The Promise and Limits of Mobile Phone Data. Journal of Infectious Diseases. 2016;214(suppl_4):S414-S20.
22. Silver L. Smartphone Ownership Is Growing Rapidly Around the World, but Not Always Equally. 2019.
23. Bol N, Helberger N, Weert JCM. Differences in mobile health app use: A source of new digital inequalities? The Information Society. 2018;34(3):183-93.
24. Hussein MR, Rahman MA, Mojumder MJH, Ahmed S, Isha SN, Akter S, et al. Trust Concerns in Health Apps collecting Personally Identifiable Information during COVID-19-like Zoonosis. 2020.
25. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 2009;6(7):e1000097.
26. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev. 2016;5(1):210.
27. Reporting guidelines 2021 [Available from: https://www.biomedcentral.com/getpublished/writing-resources/reporting-guidelines.
28. Sofaer N, Strech D. The need for systematic reviews of reasons. Bioethics. 2012;26(6):315-28.
29. Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M, et al. Guidance on the conduct of narrative synthesis in systematic reviews. A product from the ESRC methods programme Version. 2006;1:b92.
30. Farnham A, Blanke U, Stone E, Puhan MA, Hatz C. Travel medicine and mHealth technology: a study using smartphones to collect health data during travel. J Travel Med. 2016;23(6).
31. Du L, Raposo VL, Wang M. COVID-19 Contact Tracing Apps: A Technologic Tower of Babel and the Gap for International Pandemic Control. JMIR MHealth and UHealth. 2020;8(11):e23194.
32. Subramaniyaswamy V, Manogaran G, Logesh R, Vijayakumar V, Chilamkurti N, Malathi D, et al. An ontology-driven personalized food recommendation in IoT-based healthcare system. Journal of Supercomputing. 2019;75(6):3184-216.
33. Sethia D, Gupta D, Saran H. Smart health record management with secure NFC-enabled mobile devices. Smart Health. 2019;13.
34. Farnham A, Röösli M, Blanke U, Stone E, Hatz C, Puhan MA. Streaming data from a smartphone application: A new approach to mapping health during travel. Travel Medicine and Infectious Disease. 2018;21:36-42.
35. World Health O. Evidence to recommendations: methods used for assessing health equity and human rights considerations in COVID-19 and aviation: interim guidance, 23 December 2020. Geneva: World Health Organization; 2020 2020. Contract No.: WHO/2019-nCoV/Aviation/evidence_equity/2020.1.
36. Dove ES, Townend D, Meslin EM, Bobrow M, Littler K, Nicol D, et al. RESEARCH ETHICS. Ethics review for international data-intensive research. Science. 2016;351(6280):1399-400.
37. Parliament) D-GfPRSE. How the general data protection regulation changes the rules for scientific research. Publications Office ed2019.
38. Kaye J, Terry SF, Juengst E, Coy S, Harris JR, Chalmers D, et al. Including all voices in international data-sharing governance. Hum Genomics. 2018;12(1):13.

39. Ferretti A, Ronchi E, Vayena E. From principles to practice: benchmarking government guidance on health apps. Lancet Digit Heal. 2019;1(2):e55-e7.

SECTION II

THE STATE OF ETHICAL OVERSIGHT IN HEALTH RESEARCH AND HEALTH APPS: EXISTING GAPS AND KEY REFORMS

SECTION II (A)

ETHICAL OVERSIGHT IN HEALTH RESEARCH

CHAPTER 5: Big Data, Biomedical Research, and Ethics Review: New Challenges for IRBs

Peer reviewed and published as: Ferretti, A.*, Ienca, M.*, Hurst, S., Vayena, E. (2020). Big Data, Biomedical Research, and Ethics Review: New Challenges for IRBs. Ethics & Human Research, doi: 10.1002/eahr.500065

* These authors contributed equally and share first authorship

5.1 Abstract

The increased use of big data in the medical field has shifted the way in which biomedical research is designed and carried out. The novelty of techniques and methods brought by big data research brings new challenges to institutional review boards (IRBs). Yet it is unclear if IRBs should be the responsible oversight bodies for big data research and, if so, which criteria they should use. A large but heterogenous set of ethics guidelines and normative responses have emerged to address these issues. In this study, we conducted a scoping review of soft-law documents and guidelines with the aim of assessing ongoing normative efforts that are proliferating in this domain. We also synthesize a set of recurrent guidelines that could work as a baseline to create a harmonized process for big data research ethics.

5.2 Introduction

Traditionally, human subjects research in the biomedical field engages healthy and sick people as research participants in order to test certain hypotheses about health and disease according to a well-defined study design. The study designs, such as randomized controlled trials and cohort studies, are carefully reviewed by institutional review boards (IRBs)—also known as ethical review committees (ERCs) in some countries. IRBs are committed to protecting the rights and welfare of human subjects recruited to participate in biomedical or behavioral research (including social science research) (1). This approach has long been the standard for biomedical research.

Recently, however, biomedical research has begun to pursue opportunities afforded by big data. Big data research relies on large-scale databases, multiplication of data sources, advanced storage capacity, and novel computational tools that allow for high-velocity data analytics (2). In the biomedical domain, big data trends are enabled by and allow for advances in areas such as whole genome sequencing, brain imaging, mobile health, and digital phenotyping (3). Today, a large portion of health-related research relies on big data. Big data also enables researchers to draw health insights from data sources that are not strictly medical data from wearable trackers, social media, and Internet searches, for example (4). Big data research opens new prospects to accelerate health-related research and potentially elicit breakthroughs that will benefit patients (5).

Big data has been observed to shift the way biomedical researchers design and carry out their studies (6). This research departs from the traditional research model because it is largely exploratory rather than hypothesis driven. Health-related big data research is based on the acquisition of large amounts of data from multiple and often heterogeneous sources, which are subsequently combined and mined using powerful data analytics tools. This reverse-engineered approach to health-related research allows researchers to extract features and valuable insights from large datasets, without being able to anticipate exactly what the data analysis will find.

The methodological novelty of big data research models brings new challenges and questions to IRBs, including whether they are the bodies responsible for assessing these projects, and if they are, what criteria they should use to evaluate them. Given current technologies, analytic methods, and regulations, IRBs cannot take their traditional review frameworks as given. This is because big data research models might not fit within the traditional national review policies for the protection of human subjects (for example, the Common Rule in the United States and the Human Research Act in Switzerland) and the principles stated in guidelines documents such as the Helsinki Declaration of the World Medical Association and the U.S. National Commission's Belmont Report. It was observed that the definition of “human subjects” in the Common Rule might not cover big data projects involving the processing of deidentified data (7). The Common Rule's scope, in fact, is limited to the acquisition and processing of “identifiable private information”. As a consequence, privately held, publicly available datasets such as Twitter data might be considered exempt from IRB oversight, even though it is possible to reidentify those data sources by matching them with ancillary information (8). Similarly, the European Union's research ethics legislation (9), as well as the Swiss Human Research Act (10), might not apply to research that involves anonymized data or secondary use of data for which a broad consent and an ERC approval was obtained. For instance, in Denmark, researchers can reuse genomic data previously extracted from a donated tissue sample of the National Biobank Registry for a new project without seeking ERC approval (11).

Health-related big data research also challenges IRBs in referring to existing safeguards for ethics research such as informed consent, privacy and confidentiality, and minimal risk (12). The reason for that stems from a threefold consideration.

First, individuals whose data are used in research (hereafter data subjects) are often not sufficiently informed concerning the use of their data. Particularly, researchers might not be able

to adequately inform data subjects when collecting their data stored in large repositories or when mining the data in the context of secondary data uses. In a more pragmatic sense, informed consent might be hard to obtain in big data studies due to the high number of data subjects involved. This is particularly true when consent is sought retrospectively. In cases where research is conducted on large-scale repositories, it might be hardly feasible to recontact all data subjects and inform them that the purpose of data processing has changed from the original consent agreement stipulated at the time when the repository was created.

Second, breaches in data privacy and confidentiality represent a major source of risk for health research using big data. The reason for that stems from the informational richness of large research data repositories, which makes them a primary target for actors outside the research domain. Insurers, marketing companies, and the government might require access to these data. Furthermore, health-related data repositories have often been exposed to illicit use by malevolent actors. Although these repositories are usually composed of data that do not contain personal identifiers (deidentified data), research has shown that both pseudonymized data (with which artificial identifiers are used so data subjects can be reidentified) and anonymized data (from which actual and artificial identifiers are excluded in an effort to make reidentification impossible) could be matched with publicly available information or auxiliary data to allow the reidentification of a subject (13). This reidentification risk is particularly problematic for health research, as health data constitute a highly sensitive data source.

Finally, correlations arising from health-related big data analytics can be abused by various actors for unethical purposes such as discriminating against applicants to health insurance services or jobs based on health risk indicators. These indicators include, among others, risk factors associated with genetic variants, neuroimaging biomarkers of addiction or antisocial behavior, and molecular biomarkers of chronic illness. This risk of discrimination also applies to not strictly medical data such as online behavioral information. For example, a recent study has used a big data approach to predict people's sexual orientation from their online behavior (14). This poses a risk to many people, especially in countries where nonheterosexual behavior is prohibited by law. Although the research involved the processing of seemingly innocuous data points, the findings suggest that the risk of reidentification is potentially greater than minimal risk.

Many ethical issues remain to be solved. These include whether and when big data projects using deidentified data from public databases should require IRB approval, what counts as “public data”,

what constitutes “minimal-risk” in data-driven projects, and which novel ethical safeguards, if any, are required to ensure ethical big data research.

To reduce this uncertainty, various stakeholders have issued nonbinding guidelines. The scientific community has developed best-practice guidelines and educational activities aimed at sensitizing researchers about the ethical promises and challenges of big data research. In parallel, a growing number of professional organizations are restructuring their codes of conduct and providing research ethics training to data and computer scientists. Members of the scientific community, such as the editors of the journal Nature, have encouraged policy-makers to “further support such efforts… and make them better known to researchers” and have proclaimed that “all researchers have a duty to consider the ethics of their work beyond the strict limits of law or today's regulations” (15).

Nevertheless, the proliferation of many independent responses around big data research ethics has generated uncertainty among IRBs. This fragmented landscape of responses increases the confusion and leaves IRBs with unclear normative guidance about how to tackle big data ethics issues. We monitored and evaluated these efforts to bring clarity about the plurality of perspectives emerging in this domain. To accomplish our purpose, we have conducted a scoping review of the soft-law documents and guidelines (16) concerning the ethics of health-related big data research. While previous reviews have screened the scholarly literature on this topic (17), and opinion articles have discussed their implications for IRBs (18), no other study, to our knowledge, has provided a comprehensive assessment of the emerging body of guidelines on this topic. Research best practices, recommendations, codes of conduct, and other guidance documents—especially those commissioned by funding agencies, professional associations, academic societies, nongovernmental organizations (NGOs), think tanks, and private companies—are typically not formally published in peer-reviewed academic journals but, rather, released in commissioned technical reports, white papers, and similar documents. Because these types of documents are usually not indexed in academic archives and databases, reviewing this gray literature (19) is critical to retrieve and assess this body of information. We believe that IRBs will benefit from our research, as we provide a comprehensive set of recommendations that could represent the starting point for IRBs in revising and harmonizing their ethics research processes.

5.3 Approach to identifying documents for review

In February 2019, we conducted an online search of the gray literature addressing the ethical implications of health-related big data research. Gray literature is defined as “literature that is not formally published in sources such as books or journal articles” (20). This definition includes nonconventional material such as reports, technical specifications and standards, technical and commercial documentation, official documents (21), and “that which is produced on all levels of government, academics, business and industry in print and electronic formats, but which is not controlled by commercial publishing interests and where publishing is not the primary activity of the organisation” (22). Gray literature reviews have been observed to have the threefold advantage of providing information of process and implementation, both of which can be missing from scientific papers (23); reducing the typical publication lag of peer-reviewed articles, hence ensuring more efficient responses (24); and validating the results of research-based literature searches (25).

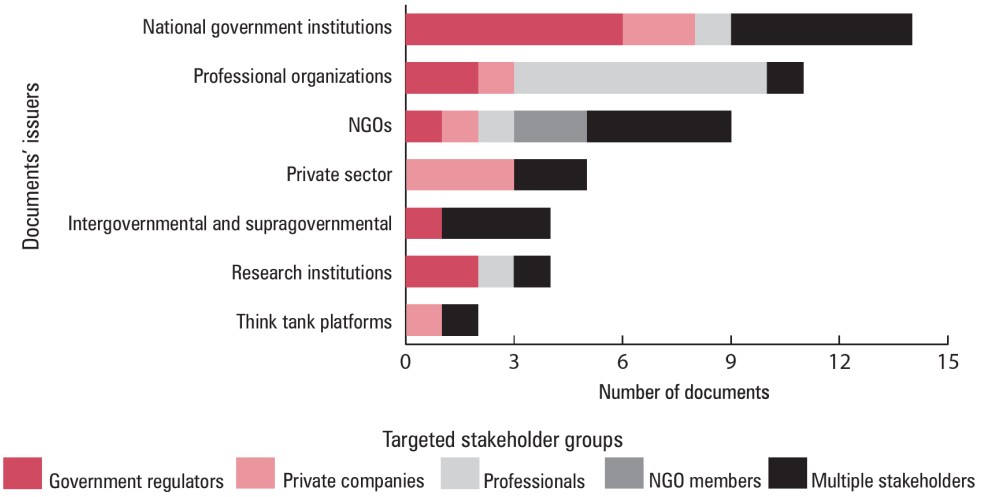
Following previous studies (26), we used a multistage screening process involving both inductive screening via search engine and deductive identification of relevant agencies (for example, national, international, and intergovernmental organizations), and subsequently we screened their websites and online collections. A total of 49 documents were included in our analysis (see Appendix 1 Chapter 5, available online, along with all the figures and the appendices; see “Supporting Information” at the end of this article). In the literature retrieval phase, we searched the Google engine in nonpersonalized mode using multiple combinations of the following keywords: “big data”, “data science”, “digital data”, “medical”, “healthcare”, “clinical”, “policy”, “ethics”, “governance”, “ethics committee”, “IRB”, and “ethics review board”. Combinations of keywords were reiterated until saturation was achieved. In the screening phase, we selected, on the one hand, soft-law documents issued by national and international agencies (highlighted in blue in Appendix 1 Chapter 5) and, on the other hand, nonlegal guidelines providing best practices and recommendations, disseminated by NGOs, professional organizations, research bodies, private companies, think tanks, and other actors (highlighted in yellow in Appendix 1 Chapter 5) (27). We included only official documents representative of and issued by collective entities. Documents such as personal blogs, written and issued by individual authors offering their personal views, were not included. In the filtering phase, we excluded from the analysis all documents that did not meet our content-based inclusion criteria (see Appendix 2 Chapter 5). The documents were collected independently by two coauthors who compared their results and

resolved interpretative discrepancies upon discussion. Documents written in English, Italian, French, Greek, and German (languages spoken by research team members) were included in the analysis. We then conducted a descriptive numerical summary and a thematic analysis. The descriptive numerical summary consisted of calculating the frequencies of the total number of articles included, the distribution of documents by the documents' issuers and the targeted stakeholder group, and the prevalence of documents with a particular health-related or IRB focus. In the latter analysis, two researchers inductively identified recurrent themes (28) with software assistance. The two researchers coded the themes using the NVivo software for qualitative data analysis (version 12 for Mac) considering three macro areas: general ethical issues, normative ethical recommendations, and specific recommendations for IRBs. Within these areas, we grouped and merged our coding in themes and macrothemes (29). Disagreement about where to allocate codes that did not seem to follow in any existing theme was resolved among coauthors through internal consultation.

5.4 Literature review results

Most documents were issued by national and governmental institutions (28%), followed by professional organizations (22%) and NGOs (20%). Fewer documents (no more than five) were issued by private companies, international institutions, research institutions, and think-tank platforms. The geographical provenance of the issuers (41% from North America, 35% from Europe, 20% being international ("international" meaning that the issuer of that document was not located in a specific country; rather, it was a multicountry agency), and 4% from Oceania) showed a higher representation of highly industrialized Western countries and a relative underrepresentation of low- and middle-income countries from the global south. The majority of the documents (35%) targeted readers from various stakeholder groups (for instance, government institutions, researchers, professionals, industry associations, consumer advocates, and the general public), with a quarter specifically targeting governmental regulations and a fifth addressing professional groups (e.g., the professional organization of statisticians) (see Figure 5.1).

Figure 5.1: Distribution of Documents by Documents' Issuers and Targeted Stakeholder Groups



As to the content of the documents, 55% had a prominent focus on health-related big data, with the remaining 45% having a broader focus on big data research, including health-research themes. Additional analysis revealed that only 16 documents (33%) were addressed explicitly to IRBs or ERCs and provided ad hoc recommendations for the review of big data projects. The remaining documents provided general recommendations concerning the general ELSI (ethical, legal, and social implications) of big data research, but they do not address IRBs directly.

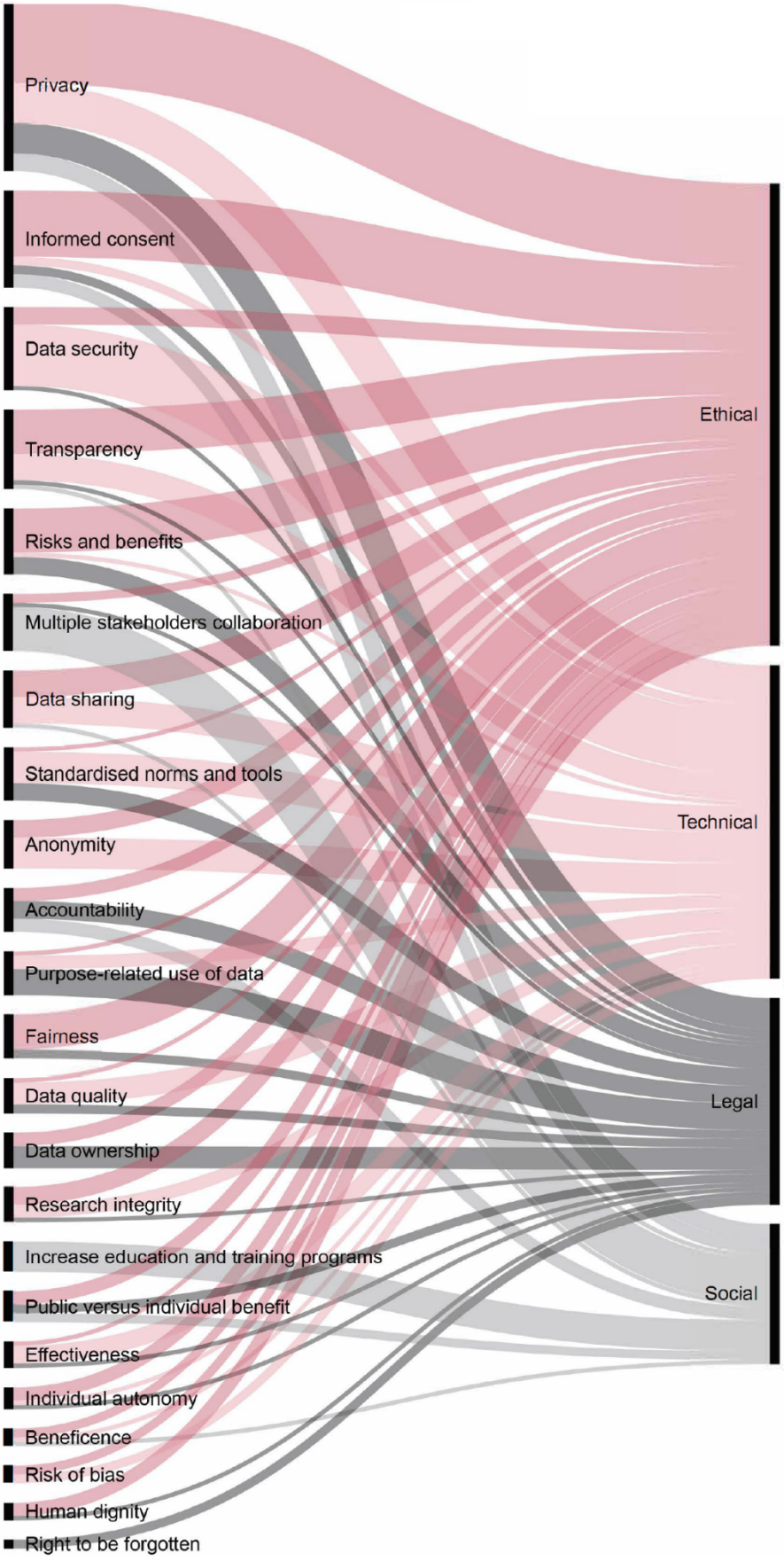
Our inductive thematic analysis identified a number of mutually interconnected ethical themes (see Figure 5.2). The notion of privacy was by far the most prevalent; it was mentioned in all documents and discussed in depth in three quarters of them. A highly recurrent issue concerned how to balance data providers' privacy while enabling the progress of research using big data techniques. While the importance of preserving privacy was widely recognized across the literature we reviewed, some documents raised the problem of harmonizing privacy regulations: given that privacy regulations differ across different jurisdictions, it might not, documents observed, be straightforward for researchers and IRBs to determine how to ensure privacy in cross-national big data research. Documents also addressed the problem of ensuring the semantic unambiguity of the privacy concept in spite of the blurred distinction between public and private data repositories in the digital ecosystem. For example, documents questioned whether health-related big data projects that use data from public-by-default social media platforms such

as Twitter should undergo similar privacy impact assessments and ethics review as conventional biomedical research does.

The second most common ethical theme was informed consent (discussed in 44% of documents), whose ethical sensitivity was primarily associated with the problem of obtaining retrospective consent from data subjects when conducting large-scale big data studies. Documents questioned the practical feasibility of retrospectively contacting many hundreds of thousands of data subjects—such as during the emotional "contagion" study (30)—and discussed the ethical justification for seeking retrospective consent in the context of public health research conducted in the public interest (such as epidemic prevention). Data ethics issues associated with data management, such as data security (39%), data sharing (37%), and data transparency (40%), also composed a significant portion of the current ethical landscape. Issues of algorithmic bias, beneficence, the right to be forgotten, data ownership, and individual autonomy appeared less prevalent.

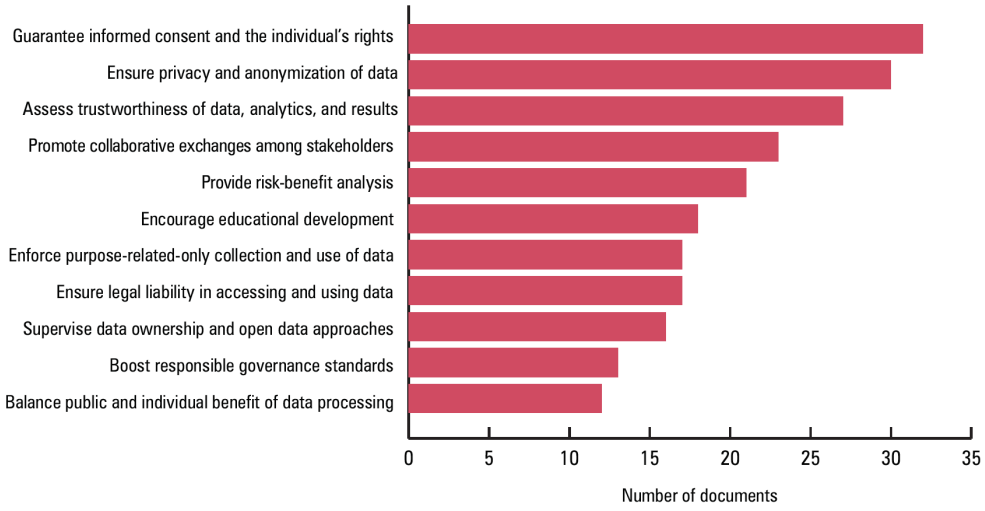
Contextual analysis of emergent themes identified the ethical, legal, social, or technical contexts within which each subtheme was discussed or a solution was proposed (see Figure 5.2). Results showed that the documents discussed 41% of the themes in the context of normative ethical analysis and with respect to potential solutions relying on ethical guidelines, best practices, or conceptual clarification. A third of the themes (31%) were discussed with reference to their technical implications and solutions. For example, distributed-ledger computing (or blockchain), encryption, and differential privacy were all mentioned as possible technical solutions to privacy risks in the big data domain (31). A smaller number of documents addressed the political and regulatory domain and proposed solutions in terms of novel legislation (18%) or social strategies (12%).

Figure 5.2: Alluvial Diagram Showing the Interconnectedness across Multiple Themes and Contextual Thematic Families



Analysis of thematic interrelations indicates a high degree of interconnectedness among (sub)themes and contexts. Although primarily discussed in the context of ethical analysis, privacy and informed consent issues were largely cross-discussed in various contexts including the technical and legislative domains. Issues of data security and stakeholder collaboration were primarily discussed in connection with, respectively, technical solutions and social considerations. Data security was primarily presented as a problem requiring technical solutions (for example, enhanced encryption, immunization to abusive apps, and adherence to international standards such as the ISO/IEC 27001:2013 (32)), whereas hard law was considered unavoidable to address issues of data ownership and to enforce the data subjects' right to be forgotten. Among these regulatory solutions, the European Union's General Data Protection Regulation, implemented in May 2018, was perceived as setting a new standard for data subject rights. Our thematic analysis also retrieved substantive recommendations for the ethics review of big data research. These recommendations concerned a variety of thematic families (see Figure 5.3) and presented notable divergences in terms of content and degree of specification.

Figure 5.3: Substantive Thematic Families of Recommendations in Big Data Research



From the perspective of substantive ethical content, the most recurrent recommendation was for IRBs to ensure that researchers are providing adequate information to data subjects as part of the process of obtaining their informed consent (n = 31, 63%). Further recommendations required data controllers and processers to protect the privacy of data subjects (n = 30, 61%) and to assure data transparency and the trustworthiness of data-driven inferences (n = 26, 53%). A smaller portion of recommendations required oversight bodies and regulators to foster collaborative

exchange among stakeholders about data uses (n = 23, 47%) and to provide new guidance to assess the benefits and harms of big data research (n = 21, 42%). Appendix 3 Chapter 5 provides further information concerning the substantive ethical recommendations issued in different continents and by various stakeholders.

In terms of granularity or degree of specification, most recommendations made general normative statements about the importance of promoting certain ethical principles (such as privacy) without specifying who should promote those principles and how, or in which domain they should be promoted and for what reason. A smaller number of documents offered a list of specifications including domain-specific and stakeholder-specific sets of good practices. Among those, a small portion of documents provided explicit recommendations for IRBs or analogous ethics review oversight bodies. In-depth thematic analysis of this subset of documents revealed four major procedural recommendations for IRBs (see Appendix 4 Chapter 5).

Strengthen oversight function

IRB oversight should be required for big data research even when the research project does not involve the physical (offline) recruitment of human subjects and does not process personally identifiable data or entail direct foreseeable harms to data generators. Furthermore, the IRB's purview should be expanded to monitor the ethical soundness of big data projects throughout the whole data lifecycle. The IRB's control mechanisms should be able to audit each phase of the project, including research planning, data collection, analytics, and results dissemination. The IRBs must inspect if ethical safeguards are in place to protect individual and group-level privacy, autonomy, safety, and the quality and transparency of data management. A few documents argued that IRBs should have the capacity to anticipate or prospectively identify if violations of data access rights might occur and to manage ethical risks associated with data disclosure—particularly when analytic techniques are in use that can allow data processors (or third parties) to reidentify individuals and reveal sensitive information. Given the expertise and independent role of IRBs, they were generally claimed to be well-suited to guarantee an impartial and objective oversight of big data studies.

Improve the review process

IRBs should improve their review process to account for the novel ethical challenges of big data studies. For example, documents noted that IRBs might need to reconsider how to review informed consent procedures in large-scale data-driven projects where traditional informed

consent models might be unfeasible (especially in the case of secondary or tertiary data uses) (33). Similarly, documents highlighted that balancing risks and benefits is increasingly complicated in the age of big data, as indirect and informational risks are harder to detect compared to the conventional physical risks of clinical research (34). Furthermore, risks such as personal data leakages across multiple data cycles are difficult to anticipate prior to data collection and during the ethics review phase. Documents highlighted the need "to create or expand accountable data ethics review processes" (35) and/or develop a new ethical framework specific to big data research. Some documents also suggested the creation of new independent advisory boards whose function should be complementary to that of IRBs (36).

Build capacity and expand competencies

Documents proposed that IRB members should receive additional training in data science and expand their knowledge of the ethical challenges of big data research. Wherever necessary, IRBs should also consider diversifying their membership to include data scientists and data ethicists. Several documents noted that IRBs are often composed of stakeholders (such as lawyers, physicians, nurses, and laypeople) who rarely have received formal training in computer or data science. Building capacity and expanding competencies are critical to anticipate and promptly identify ethical challenges. Some documents hypothesized that doing so will thereby improve the credibility of IRBs from the point of view of researchers and will increase researchers' willingness to undergo ethics review (37).

Engage with researchers

IRBs should engage more with researchers and involve them in the ethical evaluation of big data projects. The analyzed documents reported that to achieve this goal, IRBs should have an open dialogue with researchers, sensitize them about the importance of ethically aligned research, and develop facilitated channels for the ethics review of big data projects. At the same time, IRBs should get involved, together with academic ethicists, in the research ethics training of young data scientists. IRBs and researchers should establish a tight collaboration to identify, preempt, and manage ethical risks emerging in health-related big data research. Overall, we recognized high heterogeneity in the way recommendations were carried out by different issuers (see Appendix 4 Chapter 5).

We note here several study limitations. In the phase of literature retrieval, three typical limitations of gray literature reviews applied: selection bias, the volatile structure of web content, and the

documents' heterogeneity. As our search string was written in English and our inclusion criteria included only articles written in any of the languages known by one or more of the researchers, articles written in any other language have not been included. While this limitation is inherent to any literature review, we believe we have minimized it by including articles written in five languages (English, French, German, Greek, and Italian). Another possible source of selection bias is that the Google search engine results are usually returned customized to a specific user and ranked following the number of hits a website received. To anticipate this problem, the search was performed independently by two researchers using separate terminals and IP addresses and in nonpersonalized mode. Google pages were screened until data saturation was reached and acknowledged by both researchers. To minimize subjective bias, the two researchers compared their results and, in cases of divergence, discussed them with motivations until agreement was reached.

The absence of an exhaustive repository of soft-law and policy documents, together with the volatility of their web-based content, might have also affected the review. On the one hand, relevant documents might have gone undetected due to the sensitivity of our search. On the other hand, retrieved documents might, in the future, be removed from the Internet. To minimize the first risk, two coauthors independently screened multiple Google results pages and then crosschecked their results. To address the latter, we retained the original documents in PDF format and created a private repository, which will be shared upon request.

Since we included in the analysis documents that were very diverse in format, content, and quality, this heterogeneity might have affected our thematic analysis. While we are aware of this limitation, we believe that more selective inclusion criteria would have defeated the exploratory purpose of our review. Finally, inductive thematic analyses are also vulnerable to the problem of subjective interpretability by different researchers. This subjective bias is due to the methodological freedom in constructing themes by grouping codes inductively derived from the texts. Although other researchers might have chosen different classification systems, we assessed our thematic classifications iteratively and adapted them along the way to verify their consistency and adherence to the data.

5.5 Discussion

The literature review we conducted illustrates a growing corpus of soft-law documents on the ethics review of health-related big data science. The overall number of documents published on this topic increased linearly from year 2012 onwards, indicating a growing interest among regulators and other stakeholders. At the same time, the heterogeneous corpus of documents is indicative of a fragmented ethical and regulatory landscape rather than of an internationally shared framework for ethically aligned big data research. The spectrum of actors involved in this domain is diverse, as it includes, among others, regional (such as the Information and Privacy Commissioner of Ontario), governmental (such as the U. S. Office for Human Research Protections), intergovernmental (such as UNESCO), and supranational (such as the Council of Europe and the European Commission) institutions as well as private companies and NGOs. Very few documents were issued by academic research institutions, despite their direct involvement in research. In contrast, a considerable number of documents were issued by professional associations such as the United Kingdom's Royal Statistical Society and the Internet Association of Privacy Professionals. An even smaller portion of the corpus is represented by independent ethics bodies such as the Nuffield Council on Bioethics and the Italian National Bioethics Committee. However, their documents appeared ethically richer and more detailed compared to the average—an observation that is corroborated by the higher-than-average number of codes identified among these documents.

While it cannot be ruled out that private actors' involvement in big data ethics is indicative of a genuine ethical interest, it has been observed that their proactive guidance efforts have scarce democratic accountability and might raise a risk of undue influence on policy-making, especially when applied to pervasive systems such as data analytics and artificial intelligence (38). In fact, many large health-related datasets are exclusive property of companies, whose data handling and operational strategies are often hidden by nondisclosure agreements. Given the critical role of private corporations in the data economy, this industry mobilization is necessary to shape an enforceable ethical framework for big data research. At the same time, it raises the quandary of social accountability and the risk that nonstate actors might acquire a quasilegislative power. These problems have particular significance when procedural or substantive conflicts arise between the recommendations provided by, respectively, industry actors and governmental bodies.

Content analysis reveals that most documents provide general normative recommendations about the ethical, legal, and social implications of big data without specifying to which domain these recommendations apply. Moreover, from these documents, it is not clear which actors or bodies should be entitled to promote or enforce these recommendations. Only a minority of documents (33%) specifically addressed IRBs or other ethics review bodies by developing ad hoc recommendations for the review of big data projects. The reason for that is possibly twofold. First, issuer groups such as professional associations, NGOs, and private companies rarely engage with IRBs. Second, big data studies that do not involve human subjects are often perceived as falling outside the purview of ethics review (39). This interpretation is corroborated by documents such as the Menlo Report (40) and the Data & Society Report (41), which reveal that researchers involved in data-intensive research typically avoid formal ethics review, as they do not perceive it to be “human subjects research”, especially when they rely on secondary deidentified data collections or on corporate-owned databases. This result is consistent with previous studies (42) showing that researchers using big data methods are more likely to bypass IRB review and to adopt self-assessment. While self-regulation approaches are well-suited to ensure scientific freedom, bypassing ethics review via self-assessment is ethically problematic. As the history of biomedical ethics has repeatedly shown, the avoidance of independent ethics review can lead to individual or societal harm and diminish the public's trust in science (43). This is particularly true for novel areas of science whose ethical boundaries and long-term consequences are still subject to predictive uncertainty.

Both public and private actors focused their recommendations on defining the conditions for ethically sound acquisition, processing, and storage of data. The remarkable frequency of codes related to privacy and informed consent indicates a prominent ethical and practical concern around these themes. Nevertheless, ethics of big data should not be reduced solely to a privacy issue (44). Previous research has observed that, although privacy is a fundamental topic in big data research, it has been overemphasized to the detriment of other issues (45). Our findings seem to confirm this observation. Our results also indicate that ethical issues of fairness and data ownership are rarely addressed in current guidance documents. This is concerning given the largely reported risks of bias, discrimination, and informational disenfranchisement associated with algorithms and big data analytics (46). Our results are consistent with previous studies about the governance of artificial intelligence (AI) technologies that found interpretative differences and a lack of actionable requirements for the promotion of fairness and justice in the use of these technologies (47). These results indicate that attention is missing in this still-developing area of

research ethics, and they attest to persistent uncertainty on how fairness and justice considerations should be addressed in the age of big data and AI. We argue that more detailed normative guidance is needed in this regard.

The high level of interconnectedness among different ethical macrothemes highlights that ethical issues are not in silos but are intimately intertwined. This makes the ethics review of big data projects a complex and multifaceted process that involves not only scrutinizing ethical codes and methods but also inspecting technical requirements, addressing epistemological considerations, and anticipating societal implications. Results suggest that IRBs should exercise their role of essential control systems evaluating and balancing the different faces of each issue, which might require expanded purview and diversified expertise.

Given the fragmentation and heterogeneity of the current landscape of guidance documents, it is unlikely to reduce uncertainty among researchers and IRBs regarding the ethics review of health-related big data studies. Nonetheless, our results revealed a recurrence of four major procedural recommendations for IRBs. These recommendations address how IRBs should improve their review activities, strengthen their competencies, and revise some of their established practices.

First, documents identified a need for more comprehensive oversight strategies, especially by expanding the purview of IRBs to require formal ethical assessment of data-intensive studies even when they do not involve the recruitment of human subjects or operate on publicly available data repositories. At the same time, researchers should interpret ethics review not as a waiver of their responsibility but, rather, as an essential quality control of their research. While expanding the purview of IRBs might require new legislation, encouraging researchers to undergo an ethics review on a voluntary basis might be a temporary measure to improve ethical safeguards. Voluntary submissions for review can be incentivized through awareness-sensitive campaigns about the ethical implications of big data among researchers and by fast-track review procedures for projects that ensure certain technical requirements.

Second, documents urged IRBs, research institutions, and science regulators to improve the ethics review process and formalize a coherent ethical review framework for the evaluation of big data projects. Documents observed that research ethics paradigms developed for offline research are hardly transferable to data-driven research in absence of calibration. For example, research aimed at mining health-related data might have challenging implications for conceptual

milestones of human research ethics such as the notion of minimal risk. Unlike conventional research involving human subjects, big data research involving human-related data might not pose direct risks for the physical integrity of research participants. However, the last few years have borne out the fact that poorly managed datasets can have harmful consequences for human subjects in terms of mental well-being, harm to reputation, unfair treatment, and discrimination or other forms of informational risk and dignitary harm (48). Consequently, the standards of minimal risks developed for clinical research are hardly applicable to the big data domain if significant conceptual and normative adjustments are not performed.

Third, documents highlight the importance of empowering IRBs with the relevant expertise to account for the computational and ethical complexity of big data studies. IRB members trained in medicine, psychology, law, or traditional research ethics might lack the relevant expertise to determine whether, for instance, a certain project is deploying safeguards to avoid algorithmic discrimination, if the machine learning models used for decision-making are amenable to ex ante and post hoc inspection, or if group-level privacy risks can arise from the combination of differently structured data sources. Documents suggest that this epistemological gap can be filled with a two-pronged approach: by diversifying the IRB's composition and through capacity-building strategies. To diversify their composition, IRBs should consider appointing individuals with expertise in computer science, data analytics, statistics, and data ethics. Furthermore, they should consider the organization of training programs or other educational and capacity-building activities.

Expanding the IRB purview and their members' expertise is a requirement grounded on the assumption that IRBs should be the relevant oversight body of big data research. This assumption was not shared unanimously. A few documents addressed the issue of whether IRBs should be the oversight body accountable for big data research at all (49). For example, data protection officers were proposed as complementary oversight resources. The creation of novel oversight bodies such as data boards was also proposed as an adaptive governance solution to the big data ethics conundrum.

Finally, documents highlighted the importance of sensitizing researchers and other relevant actors (for instance, technology developers, data analysts, advertisers, insurers, and physicians) about data ethics. The persistent absence of an agreed-upon ethical framework for big data research might perpetuate uncertainty between both researchers and IRBs and could result in

divergent approval decisions. Raising awareness within the research community can help reduce this uncertainty through proactive measures such as the development of codes of ethics and professional conduct (as done by the Association for Computing Machinery and the British Computer Science Association), research roadmaps (as done by the Association of the British Pharmaceutical Industry), or best practices (as done by the Health IT Policy Committee of the United States). Any development of an ethical framework for big data research, however, cannot disregard the active involvement of IRBs in decision-making. On the contrary, research on the views, needs, and attitudes of IRB members is highly necessary to set an evidence-based, empirically informed agenda for big data and research ethics.

Despite the prevalence of the above-listed recurrent themes across documents, there is still much uncertainty about how the recommendations should be implemented. For instance, it is not clear yet how, in practice, IRBs should improve the ethics review process and which recommendation should be implemented first. Whether IRBs should be the bodies devoted to assessing big data projects at all is still debatable. Alternatives might involve universities providing ethical requirements to their researchers who collect, store, or use big data. Additionally, peer-reviewed journals might set the rule to reject all those publications that do not follow specific ethical procedures and criteria. Another option could involve producing new legislation that includes new research ethics best practices. When advancing this ethical discussion, it is critical that IRBs are not considered passive recipients of guidelines but are actively involved in the norm-development process. To achieve this aim, qualitative studies assessing the views, needs, and attitudes of IRBs as well as collaborative approaches to guideline development are highly needed.

Acknowledgements

Agata Ferretti and Marcello Ienca contributed equally to this article. This research was supported by the Swiss National Science Foundation under award 407540_167223 (NRP 75 Big Data).

References

1. Hershey, N., “IRB Jurisdiction and Limits on IRB Actions”, IRB: Ethics & Human Research 7, no. 2 (1985): 7– 9; Cook, A. F., and H. Hoas, “Protecting Research Subjects: IRBs in a Changing Research Landscape”, IRB: Ethics & Human Research 33, no. 2 (2011): 14-19.
2. Ta, V. D., C. M. Liu, and G. W. Nkabinde, “Big Data Stream Computing in Healthcare Real-Time Analytics”, in 2016 IEEE International Conference on Cloud Computing and Big Data Analysis (Chengdu: IEEE, 2016): 37– 42.
3. Marx, V., “The Big Challenges of Big Data”, Nature 498, no. 7453 (2013): 255– 60; Greely, H. T., K. M. Ramos, and C. Grady, “Neuroethics in the Age of Brain Projects”, Neuron 92, no. 3 (2016): 637-41; Landhuis, E., “Neuroscience: Big Brain, Big Data”, Nature 541 (2017): 559-61; Insel, T. R., “Digital Phenotyping: Technology for a New Science of Behavior”, Journal of the American Medical Association 318 (2017): 1215-16.
4. Vayena, E., and U. Gasser, “Strictly Biomedical? Sketching the Ethics of the Big Data Ecosystem in Biomedicine”, in The Ethics of Biomedical Big Data (Cham, Switzerland: Springer, 2016): 17– 39.
5. Ienca, M., E. Vayena, and A. Blasimme, “Big Data and Dementia: Charting the Route Ahead for Research, Ethics, and Policy”, Frontiers in Medicine 5, no. 13 (2018): doi.org/10.3389/fmed.2018.00013.
6. Mazzocchi, F., “Could Big Data Be the End of Theory in Science? A Few Remarks on the Epistemology of Data-Driven Science”, EMBO Reports 16, no. 10 (2015): 1250– 55.
7. Metcalf, J., and K. Crawford, “Where Are Human Subjects in Big Data Research? The Emerging Ethics Divide”, Big Data & Society 3, no. 1 (2016): doi:10.1177/2053951716650211.
8. Vayena, E., et al., “Elements of a New Ethical Framework for Big Data Research”, Washington and Lee Law Review Online (2016): doi.org/10.3389/fmed.2018.00013.
9. European Commision, Ethics for Researchers: Facilitating Research Excellence in FP7 (Brussels: European Commission, 2013), http://ec.europa.eu/research/participants/data/ref/fp7/89888/ethics-for-researchers_en.pdf.
10. SAMS, “Research with Human Subjects”, Swiss Academy of Medical Sciences, (2015), https://www.sams.ch/dam/jcr:33181182-5ed6-4004-addc-86535089cfd9/handbook_sams_research_2015.pdf.
11. Holm, S., and T. Ploug, “Big Data and Health Research—the Governance Challenges in a Mixed Data Economy”, Journal of Bioethical Inquiry 14, no. 4 (2017): 515– 25.
12. Ienca, M., et al., “Considerations for Ethics Review of Big Data Health Research: A Scoping Review”, PloS One 13, no. 10 (2018): doi.org/10.1371/journal.pone.0204937; Vayena, E., and A. Blasimme, “Health Research with Big Data: Time for Systemic Oversight”, Journal of Law, Medicine & Ethics 46, no. 1 (2018): 119-29; Christen, M., et al., “On the Compatibility of Big Data Driven Research and Informed Consent: The Example of the Human Brain Project”, in The Ethics of Biomedical Big Data, ed. Mittelstadt, B. D., L. Floridi (Cham, Switzerland: Springer, 2016): 199-218; Ioannidis, J. P., “Informed Consent, Big Data, and the Oxymoron of Research That Is Not Research”, American Journal of Bioethics 13, no. 4 (2013): 40-42; Knoppers, B. M., and A. M. Thorogood, “Ethics and Big Data in Health”, Current Opinion in Systems Biology 4 (2017): 53-57.
13. El Emam, K., S. Rodgers, and B. Malin, “Anonymising and Sharing Individual Patient Data”, BMJ 350 (2015): h1139.
14. Wang, Y., and M. Kosinski, “Deep Neural Networks Are More Accurate Than Humans at Detecting Sexual Orientation from Facial Images”, Journal of Personality and Social Psychology 114, no. 2 (2018): 246.

15. “Cambridge Analytica Controversy Must Spur Researchers to Update Data Ethics”, editorial, Nature 555 (2018): doi:10.1038/d41586-018-03856-4.
16. EurWORK (European Observatory of Working Life), “Soft Law”, Eurofound, May 4, 2011, https://www.eurofound.europa.eu/observatories/eurwork/industrial-relations-dictionary/soft-law; Sossin, L., and C. W. Smith, “Hard Choices and Soft Law: Ethical Codes, Policy Guidelines and the Role of the Courts in Regulating Government”, Alberta Law Review 40, no. 4 (2003): 867– 93.
17. Ienca et al., “Considerations for Ethics Review.”
18. Dereli, T., et al., “Big Data and Ethics Review for Health Systems Research in LMICs: Understanding Risk, Uncertainty and Ignorance—and Catching the Black Swans?”, American Journal of Bioethics 14, no. 2 (2014): 48– 50.
19. McGrath, Y., et al., Review of Grey Literature on Drug Prevention among Young People (London: NICE, 2006); Favin, M., et al., “Why Children Are Not Vaccinated: A Review of the Grey Literature”, International Health 4, no. 4 (2012): 229-38.
20. Higgins, J. P., and Green, S., “Cochrane Handbook for Systematic Reviews of Interventions”, (2008), https://training.cochrane.org/handbook.
21. Alberani, V., P. D. C. Pietrangeli, and Mazza, A., “The Use of Grey Literature in Health Sciences: A Preliminary Survey”, Bulletin of the Medical Library Association 78, no. 4 (1990): 358.
22. Aina, L., “Grey Literature and Library and Information Studies: A Global Perspective”, International Journal on Grey Literature 1, no. 4 (2000): 179– 82.
23. McGrath et al., Review of Grey Literature.
24. Daniulaityte, R., et al., “Qualitative Epidemiologic Methods Can Improve Local Prevention Programming among Adolescents”, Journal of Alcohol and Drug Education 48 (2004): 73– 84.
25. Benzies, K. M., et al., “State-of-the-Evidence Reviews: Advantages and Challenges of Including Grey Literature”, Worldviews on Evidence-Based Nursing 3, no. 2 (2006): 55–61.
26. McGrath et al., Review of Grey Literature; Lawrence, A., “Influence Seekers: The Production of Grey Literature for Policy and Practice”, Information Services & Use 37, no. 4 (2017): 389– 403; Simkhada, P., et al., “Chasing the Grey Evidence: A Standardised Systematic Critical Literature Review Approach”, paper presented at the GL-conference series, 2005, http://www.opengrey.eu/item/display/10068/697851.
27. EurWORK, “Soft Law”; Sossin and Smith, “Hard Choices and Soft Law.”
28. Herzog, C., C. Handke, and E. Hitters, “Analyzing Talk and Text II: Thematic Analysis”, in The Palgrave Handbook of Methods for Media Policy Research, ed. H. Van den Bulck et al. (Basingstoke, England: Palgrave Macmillan, 2019), 385– 401.
29. Braun, V., and V. Clarke, “Using Thematic Analysis in Psychology”, Qualitative Research in Psychology 3, no. 2 (2006): 77– 101.
30. Kramer, A. D., J. E. Guillory, and J. T. Hancock, “Experimental Evidence of Massive-Scale Emotional Contagion through Social Networks”, Proceedings of the National Academy of Sciences (2014): 8788– 90.
31. Jagadish, H. V., et al., “Big Data and Its Technical Challenges”, Communications of the ACM 57, no. 7 (2014): 86– 94.
32. “ISO/IEC 27001:2013”, ISO (International Organization for Standardization), https://www.iso.org/standard/54534.html.
33. U.K. Data Service, Big Data and Data Sharing: Ethical Issues (updated February 2017), https://www.ukdataservice.ac.uk/media/604711/big-data-and-data-sharing_ethical-issues.pdf; Information Commissioner's Office, Big Data, Artificial Intelligence, Machine Learning and Data Protection (2017), https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf.

34. Council of Europe, Guidelines on the Protection of Individuals with Regard to the Processing of Personal Data in a World of Big Data (Strasbourg, France: Council of Europe, 2017), https://rm.coe.int/16806ebe7a; “Attachment A: Human Subjects Research Implications of ‘Big Data’ Studies”, Office for Human Research Protections, 2015, https://www.hhs.gov/ohrp/sachrp-committee/recommendations/2015-april-24-attachment-a/index.html.
35. Tene, O., and J. Polonetsky, “Big Data for All: Privacy and User Control in the Age of Analytics”, Northwestern Journal of Technology and Intellectual Property 239 (2013): article 1.
36. Council of Europe, Guidelines on the Protection of Individuals; International Association of Privacy Professionals, Building Ethics into Privacy Frameworks for Big Data and AI (Strasbourg, France: Council of Europe, 2017), https://rm.coe.int/16806ebe7a.
37. Boyd, D., E. F. Keller, and B. Tijerina, “Supporting Ethical Data Research: An Exploratory Study of Emerging Issues in Big Data and Technical Research”, working paper, Data & Society, August 4, 2016, https://www.datasociety.net/pubs/sedr/SupportingEthicsDataResearch_Sept2016.pdf; Nuffield Council of Bioethics, “Response to the Science and Technology Select Committee (Commons) Inquiry: The Big Data Dilemma”, September 2015, http://nuffieldbioethics.org/wp-content/uploads/Big-Data-dilemma-Nuffield-Council-on-Bioethics-September-2015.pdf.
38. Benkler, Y., “Don't Let Industry Write the Rules for AI”, Nature 569 (2019): 161.
39. “Cambridge Analytica Controversy Must Spur Researchers to Update Data Ethics”, editorial, Nature (March 27, 2018), https://www.nature.com/articles/d41586-018-03856-4.
40. The Menlo Report, 2012, https://www.caida.org/publications/papers/2012/menlo_report_actual_formatted/menlo_report_actual_formatted.pdf.
41. Supporting Ethical Data Research: An Exploratory Study of Emerging Issues in Big Data and Technical Research, 2016, https://www.datasociety.net/pubs/sedr/SupportingEthicsDataResearch_Sept2016.pdf.
42. Samuel, G. N., and B. Farsides, “Public Trust and ‘Ethics Review’ as a Commodity: The Case of Genomics England Limited and the UK's 100,000 Genomes Project”, Medicine, Health Care and Philosophy (2017): 159– 68; Metcalf and Crawford, “Where Are Human Subjects in Big Data Research?”
43. Gracia, D., “History of Medical Ethics”, in Bioethics in a European Perspective, ed. H. T. Have and B. Gordjin (Cham, Switzerland: Springer, 2001), 17– 50.
44. Mason, P. H., The Ethics of Biomedical Big Data, ed. B. D. Mittelstadt and L. Floridi (Cham, Switzerland: Springer International Publishing, 2016); Ienca et al., “Considerations for Ethics Review.”
45. Mittelstadt, B. D., and L. Floridi, “The Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts”, Science and Engineering Ethics 22, no. 2 (2016): 303– 41.
46. Cuquet, M., and A. Fensel, “The Societal Impact of Big Data: A Research Roadmap for Europe”, Technology in Society 54 (2018): 74– 86; Knoppers and Thorogood, “Ethics and Big Data in Health.”
47. Jobin, A., M. Ienca, and E. Vayena, “The Global Landscape of AI Ethics Guidelines”, Nature Machine Intelligence 1, no. 9 (2019): 389– 99.
48. “Cambridge Analytica Controversy Must Spur Researchers to Update Data Ethics”, editorial, Nature.
49. Council of Europe, Guidelines on the Protection of Individuals; International Association of Privacy Professionals, Building Ethics into Privacy Frameworks.

CHAPTER 6: The Challenges of Big Data for Research Ethics Committees: A Qualitative Swiss study

Submitted for review as: Ferretti, A., Ienca, M., Rivas Velarde, M., Hurst, S., Vayena, E. The Challenges of Big Data for Research Ethics Committees: A Qualitative Swiss study.

6.1 Abstract

Big data trends in health research challenge the oversight mechanism of Research Ethics Committees (RECs). Traditional standards of research quality and the mandate of RECs illuminate deficits in facing the computational complexity, methodological novelty, and limited auditability of these approaches. To better understand the challenges facing RECs, we explored the perspectives and attitudes of the chairpersons and scientific assistants of the seven Swiss Cantonal RECs via semi-structured qualitative interviews. Our interviews expose interviewees' minimal experience with reviewing big data research, insufficient expertise in data science, and uncertainty about how to mitigate big data research risks. Nonetheless, RECs could strengthen their oversight by training in data science and big data ethics, complementing their role with external experts and ad-hoc boards, and introducing precise shared practices.

6.2 Introduction

In recent years, research using large volumes of data has drastically increased across a variety of fields including data science, physics, biomedicine, psychology, and the social sciences (1). This type of research, known as big data research, benefits from merging and harnessing data from multiple sources, generating new insights and unexplored scientific perspectives. In parallel with these changes in research practice, high profile cases of data misuse emerged, exposing research participants to significant risk of harm (2). In response, debate has increased about the role and effectiveness of the research ethics committee (REC) as the chief ethical research oversight mechanism in research, given the specific challenges presented by research with big data (3, 4). RECs, also known as Institutional Review Boards (IRBs) and Research Ethics Boards (REBs), were created in the twentieth century to protect the safety and interests of human participants in research (5). Today, the REC's mandate – the regulation of human subject research and the evaluation of key ethics review principles – might fall behind the demands of data-intensive research (6). In fact, big data research is characterized by novel ethical concerns which can challenge traditional ethics oversight mechanisms and practices (7).

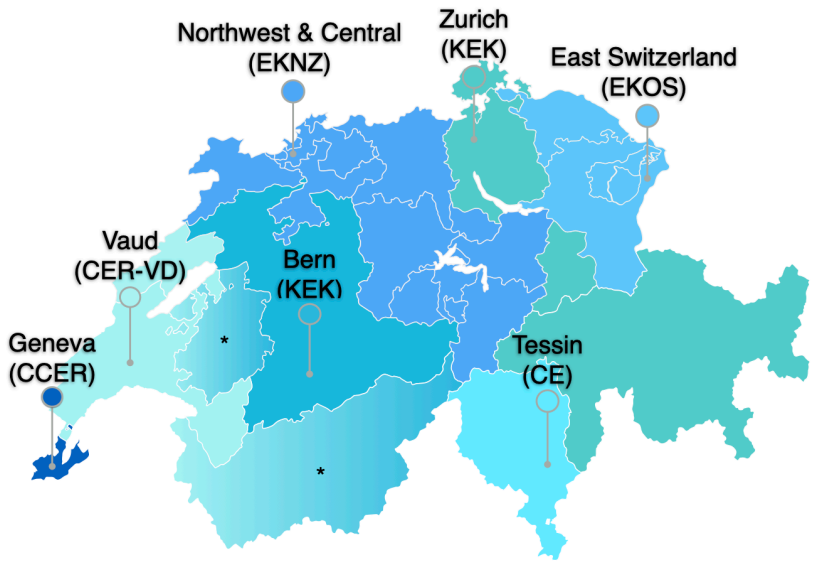
Particularly in the biomedical and health fields, the increasing availability of digital health technologies enables the collection of an unprecedented amount of data (8). In addition, the

possibility of using artificial intelligence (AI) and extraordinary computational capabilities to merge, analyze, and harness these data offers great opportunities to improve individual and public health (9). The potential of artificial intelligence in medicine has emerged even more clearly during the COVID-19 pandemic, as differently structured data from heterogenous sources were collected and processed for public health purposes, such as containment, mitigation, and vaccine development (10, 11). A crucial benefit offered by AI technologies is improved prevention and personalized treatment. In fact, AI can extract information related to individual health status by combining data unrelated to health and wellbeing (e.g., location data, metadata, blog posts) collected through a variety of tools (e.g., social media, wearable devices, mobile phones, smart devices) (12). Despite the mentioned benefits, these new research methods and technological developments have numerous downsides. First, they challenge traditional research principles such as data privacy, informed consent, scientific validity of research, risk assessment, and distribution of benefits (13, 14). Second, they introduce new epistemic challenges related to the assessment of scientific validity, technological reliability, accountability, fairness, and transparency (15). Finally, they challenge the very notion of human participants in research, as they enable retrospective data processing without physical interaction with research participants (16).

Several questions arise about whether existing regulatory and ethical governance tools, as well as the current practices and expertise of RECs, are adequate to protect human participants and enable ethical research (17). While some authors argue that the ethical principles and frameworks that traditionally govern research need to be adapted considering new research contexts (18, 19), studies that investigate the perspectives and needs of the involved stakeholders remain scarce. Recent studies (7, 20) analyzing researcher views on the topic revealed both a lack of adequate expertise among REC members and the absence of clear and consistent criteria for evaluation. Similar conclusions were reached by empirical studies conducted in the UK, Canada, and the United States involving interviews with REC members about the ethics of social media research and research using pervasive sensing technologies (21-23). In these studies, REC members were able to identify emerging ethical challenges related to big data but reported feeling unprepared to address those challenges, and a lack of normative guidance. Although these studies are highly informative, their exploratory and context-dependent nature makes their claims difficult to generalize. Furthermore, it should be noted that ethical oversight practices and research ethics guidelines diverge at the international level because legal requirements differ from state to state (24).

In Switzerland, for example, research projects involving human subjects, biomedical data, and biological samples require the approval of a REC. Most of the research projects conducted in biomedical and health fields are reviewed by Cantonal RECs (25). Switzerland counts seven of these committees which are organized under Swissethics, the Swiss association of Cantonal RECs (26). RECs apply the legal and ethical rules included in the Human Research Act (HRA), which ensure the dignity, privacy, and health of research participants, as well as the ethical value of the research. Each REC oversees projects falling under the HRA requirements in a specific geographical area of Switzerland. There are two committees in the French-speaking region of Switzerland, one in the Italian-speaking region, and four in the German-speaking region (Figure 6.1).

Figure 6.1: Distribution of Swiss RECs in the Swiss territory and areas of authority



* Covered by both Bern and Vaud RECs

Typically, research involving anonymously collected health and biomedical data or anonymized biological samples is not subject to the HRA. Similarly, studies that do not have direct implications for "the understanding of human diseases; the structure and function of the human body; or public health" (27) are exempted under this regulation. As a consequence, human subject research in the fields of psychology, sociology, or marketing, as well as projects using non-strictly biomedical data to infer the health status of individuals, are exempted from the HRA provisions. Several Swiss universities have introduced institutional ethics committees to review those research

projects that fall outside the Cantonal RECs purview. Nevertheless, the implementation of such intra-institutional local ethics committees is fragmented and uneven within the country, since universities have no legal obligation to introduce these committees. In fact, federal law only provides for the establishment of Cantonal RECs.

While a recent study looked at the experience of Swiss researchers when submitting their big data research for ethical review (28), no study to date has investigated the opinions and perspectives of Cantonal RECs. Therefore, this study aims to fill this gap and expand knowledge on the topic by engaging with members of Cantonal RECs. Their direct experience in reviewing and evaluating big data projects can provide valuable insight into the current state of the primary ethical oversight mechanism in Switzerland. In addition, their needs and attitudes can shed light on existing gaps in the mechanism, and consequently pave the way for needed reforms.

6.3 Methods

Recruitment and sampling

We conducted interviews with the chairperson (or the vice-chairperson) and, whenever possible, the scientific secretary of all Swiss Cantonal RECs. The committees were identified through the Swissethics website. The invitation email sent to participants included the following: the outline of the research and research aims; the interview methodology and a preliminary timeline; the informed consent form and details about safeguards in place for data protection and confidentiality; and the research team contacts. The response rate was 100%. All Cantonal RECs (n=7) responded to our invitation and participated in our study. Prior to recruitment, we obtained approval (EK 2017-N-74) to conduct this study from ETH Zurich's Research Ethics Committee.

Interviews

Between October 2018 and May 2019, MI, AF, and MRV conducted semi-structured interviews, either face-to-face (at the interviewees' institutions) or via telephone. Each interview was recorded, after written and verbal consent was granted, and lasted between 35 minutes and 1 hour (approximately 45 minutes on average). Interviewees could specify their preference for the interview language, from French, German, Italian, English, or a combination of these. For each

REC, we interviewed the chairperson and the scientific secretary, obtaining a total of 7 interviews with 13 interviewees (7 females and 6 males). Across RECs, interviewees shared similar disciplinary backgrounds. Most specialized in biology, medicine, and pharmacology, while only a few trained in public health, law, or statistics (Table 6.1).

MI developed the interview guide (Appendix 1 Chapter 6) which was vetted by AF and EV and approved by the entire research team. These interviews aim to investigate the perspective of the Cantonal RECs on i) how to define big data research; ii) their experience with reviewing big data projects and with the ethical guidelines used for assessment of big data research; iii) the peculiarities of big data research, namely its benefits and challenges; iv) the needs of RECS in order to adequately address big data research challenges (e.g., high-level recommendations, procedural good practices, education, training).

Table 6.1: Demographics

N. of interviews	**Total n. of interviewees**	**Gender of interviewees**		**Interviewees' fields of expertise ***					
		Female	Male	Biology	Medicine	Pharmacology	Public Health	Law/health law	Statistics
7	13	7	6	6	6	3	2	2	1

* Each interviewee can be expert in more than 1 field.

Analysis

We transcribed *verbatim* the audio files in the original language of the interviews with the support of the *Sonix* online software. Three interviews were in English, three in German, and one in Italian. To increase data consistency and reduce selective bias in the analysis, we translated the four non-English transcriptions into English with the assistance of the *DeepL Pro* online software and additional human review. AF, MI, and MRV thematically coded and analyzed the data with the *NVivo 11* Software. Each interview was coded independently by two researchers using a combination of inductive and deductive reasoning. The deductive analysis traced the themes listed in the interview guide, and was expanded by adding themes that emerged during the interviews. The data analysis was then performed in two steps. First, major themes of interest were identified and categorized (please refer to the Results section for an overview of the identified themes). This phase was duplicated by two researchers, and any disagreement was resolved through involving a third researcher. Second, the themes were analyzed in depth

through discussion among the researchers, and adjustments to the final thematic map were made to improve logical cohesion. The result of this analysis is detailed in the following section.

6.4 Results

Our analysis identified four recurrent themes and several subthemes, which are summarized in Table 6.2. The first theme relates to the RECs' understanding of the "big data" concept. The second concerns the ethics review process currently in place for big data research. The third comprehends respondents' perspectives about the benefits and challenges of big data research. Finally, the fourth emphasizes the needs of RECs in the big data era.

Table 6.2: Overview of interview themes and subthemes

	THEMES	SUBTHEMES
I.	**Characteristics of big data research**	- Variation in big data research definitions - Examples of big data research
II.	**REC mechanism in big data research**	- Frequency of big data research review - RECs' limited oversight mandate - Criteria and guidelines used to review big data research - Exceptionalism of big data research
III.	**Implications of big data research**	- Benefits - Challenges
IV.	**RECs' needs in big data research**	- Training needs - Procedural needs - Regulatory needs

Characteristics of big data research

When asked to define big data research, the interviewees displayed variation in interpreting this concept. Some of them – e.g. Interviewee 10 – observed that "*there is much talking about big data but no unanimous definition*". The majority of interviewees mentioned the three Vs (volume, variety, and velocity) characterizing big data. Particularly, they stressed the aspects of volume and variety.

> *"To me, what is relevant is data volume… the fact that there is an increasing amount of data in research files or databases. In addition, it is important where and how these data come from"* (Interviewee 10)

Instead, only a minority associated the big data concept with the deployment of novel analytic tools – such as algorithms and artificial intelligence (AI).

> "*I think that what would qualify as big data approach […] is if data are being analyzed using artificial intelligence and other analytic approaches that usually are not used for the regular project that we are evaluating - where normal or ordinary [statistical] methodology is applied. Here, with the big data, you are getting into a new dimension."* (Interviewee 12)

Furthermore, while Interviewee 13 stressed the fact that big data projects are often hypothesis-free ("*[big data projects] will try to generate the knowledge from the data itself rather than the classical approach with hypotheses and verifications*"), Interviewee 3 suggested considering data transfers and re-uses of existing datasets as signals of big data research.

Although interviewees could formulate definitions of big data research, they were confused about where to draw the line between traditional and big data research ("*when does a biomedical project start to be a big data project?*" (Interviewee 3)). Interviewee 4 said that medical research always collected and relied upon voluminous datasets. Therefore, it is only a matter of interpretation whether traditional research is considered big data research:

> "*in cancer research it is common to integrate many patients' pathology data with x-rays or other imaging data, and genetic information. This happened already in small projects; but now these projects are viewed as big data projects*" (Interviewee 4).

We asked interviewees to describe examples of big data projects they had reviewed or foresaw reviewing. Many interviewees referred to projects using data and samples stored in biobanks. Others spoke about projects focused on improving personalized medicine, using data recorded through tracking devices and wearables, as well as through social media (i.e., Facebook and Twitter) and the internet of things.

"When you talk to me about big data my idea goes more to databases, or biological sample banks that collect a huge amount of data and for which there is no purpose. […] I think big data means analyzing a huge amount of data from various sources but without a precise purpose in mind." (Interviewee 8)

"*People can be monitored at home with the goal to provide calibrated assistance. Many people are actually already completely tracked [by devices], now they lay down in beds with sensors to check how they sleep [...]."* (Interviewee 2)

Current state of ethical oversight in big data research

Six out of seven Cantonal RECs reported having previously reviewed and assessed big data projects. Nevertheless, our respondents emphasized that, so far, this had occurred rarely, only a few times a year per REC. Moreover, none of these studies were explicitly labeled by the researchers as a big data project. Interviewees acknowledged their limited experience in reviewing big data research and speculated that this is because only a few of these projects had taken place in Switzerland so far. However, REC members anticipated that this trend would evolve in the future, especially with the creation of new biobanks and more medical data collected in digital form via electronic health records (EHRs). Furthermore, interviewees highlighted that their oversight power is limited in the big data context. Their precisely defined mandate might be a reason they only rarely reviewed big data projects. In fact, Cantonal RECs' research purview is restricted to biomedical and clinical projects involving humans, and human biological data and samples. For instance, big data studies collecting social media data or anonymized data in the fields of social sciences and psychology would fall outside Cantonal RECs' review:

"They [the not-strictly biomedical projects] are, so to speak, in the grey area: the conventional ethics committees are not responsible for them, but it is still completely unclear which oversight mechanisms should be applied" (Interviewee 1)

As Interviewee 4 pointed out, Cantonal RECs may audit the above-mentioned studies, but only to *"give an opinion [not-legally binding] according to article 51 [of the Swiss Human Research Act] about whether or not these types of research applications are ethical*". As a result, this happens

only rarely because, as other REC members commented, researchers are not legally required to submit these types of projects for review. Consequently, RECs are unaware of the real state of the art concerning big data research:

> *"I am really wondering actually whether the researchers doing research on big data are willing to come and ask for our opinion. I will not be so surprised to learn that there are researchers that have actually conducted research on big data without coming to us, and I believe that under the legal point of view they may have some arguments."* (Interviewee 13)

However, those big data projects which come under the scrutiny of RECs are not treated differently than traditional biomedical research. All interviewees reported the absence of specific standards to assess big data projects. Therefore, REC members rely on traditional research ethics and bioethics criteria (such as those included in the HRA, Belmont Report (29), Emanuel framework (30), and Beauchamp's four-principle approach (31)), independent of the study type. RECs' assessment includes evaluation of data protection safeguards, strategies to respect participant autonomy (i.e., informed consent form), risk-benefit assessments, research purposes and data proportionality, and the scientific validity of both research methodology and findings:

> *"For now, there is no evaluation grid for analyzing these studies involving big data. [...] The purpose of the study, the scientific question to which the study responds, is fundamental, and is one of the factors that we take into account."* (Interviewee 8)

> *"[In big data research] I use the same usual criteria and ask 'What is the scientific question? How is it approached? Is this [question] compatible with the volume of data analyzed? What tools are used?...you know the analytic methods... Do they bring any risk?"* (Interviewee 13)

Our interviews revealed divergence among REC members' opinions about whether the lack of specific guidance for big data research should be interpreted as potentially problematic.

Interviewee 3, for example, explained that the absence of specific guidance for big data research should not necessarily be considered a weakness in the oversight mechanism. On the contrary,

existing regulations provide tools that can be effectively applied across scientific disciplines, project types, and to big data research too:

> *"I think for that what we are seeing at the moment...I think this ... we have a law, we have data protection rules, we have the Human Research Act here and I think the regulations we have can apply for this kind of research [big data research] as well as for other types of research. So, we should not make any difference at the moment."* (Interviewee 3)

A few interviewees agreed with this perspective, and spoke about the HRA, Swiss data protection law, GDPR, and Emanuel framework for biomedical research as sufficient tools to guide their judgment when reviewing projects. Two interviewees openly rejected the concept of big data research exceptionalism:

> *"I mean, of course big data shows that issues are more pressing to answer. But the pending questions...we have identified them, even though from a different point of view. [...] for each of those issues I can provide examples in traditional research that are already raising those questions."* (Interviewee 12)

> *"I don't want big data to be defined any differently than other requests […] only because it's called big data… for me it is not fundamentally different than a normal request."* (Interviewee 4)

Other interviewees, however, stated that big data research is not comparable to traditional clinical or biomedical research. These research types diverge not only in terms of data volume and sources, but also in the types of risks involved. To demonstrate this point, Interviewee 1 commented on the deployment of opaque algorithms such as deep learning to find correlations in the data, and the potential unforeseeable risks emerging from it:

> *"…if you've done this [assessing projects] for so many years now, you have a certain routine. But with big data and AI, if we don't even know what the risks are, how can we assess and approve them?"* (Interviewee 1)

In addition, Interviewee 10 spoke about the difficulty of assessing data quality in big data research compared with traditional biomedical research. If data were collected inside hospitals by

researchers and health professionals in traditional biomedical research, these data are now collected by tracking devices or social media platforms:

> *"One problem is that there is no control for data quality in self-collected self-tracked data. All those medical apps, all those devices. There is no quality control for that. Who is ensuring, checking the quality of the data they generate? Same for social media… We need a big switch in our mentality: The quality of those data is not, at least not always, identical as in conventional research."* (Interviewee 10)

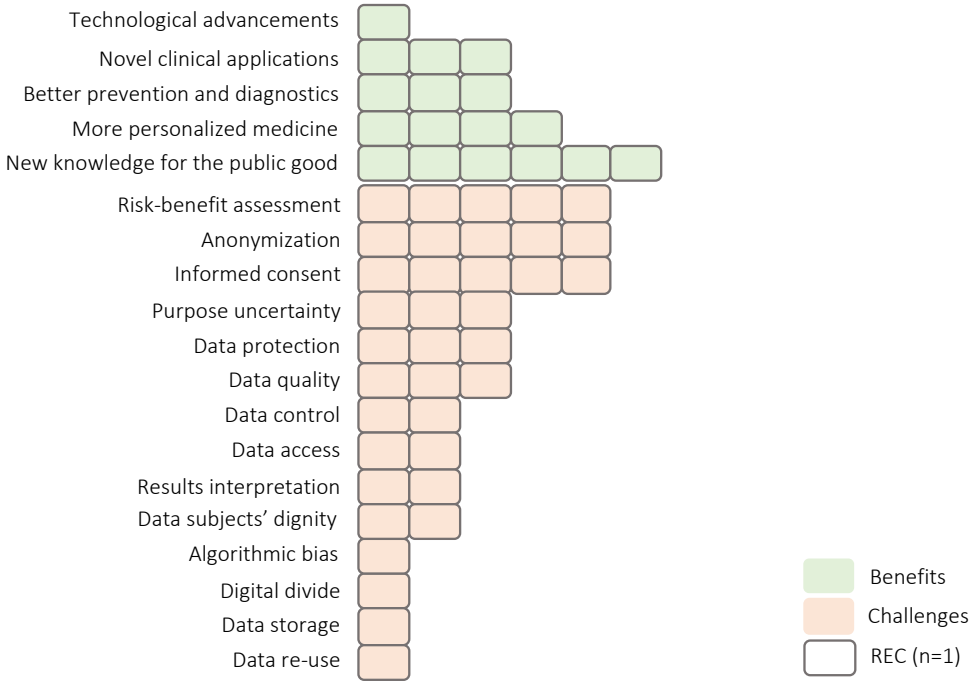
Finally, Interviewee 9 provided an example of how the absence of clear legal guidance can result in inconsistent ethical evaluations across RECs:

> *"If there is not a sufficient legal framework...well, the problem is that projects involving big data are only interpreted from an ethical point of view. And ethical interpretations from one committee to another may vary. The lack of a precise frame allows you to have more interpretations – which are always interesting – but could create problems. We would review these projects with the risk that our assessment could be called into question."* (Interviewee 9)

Implications of big data research

Overall, our respondents indicated a variety of benefits and challenges associated with big data research (these are summarized in Figure 6.2).

Figure 6.2: Benefits and challenges of big data research discussed by Swiss Cantonal RECs



Concerning the former, nearly all REC members flagged the importance of big data studies for increasing scientific knowledge and generating public benefit; "*I certainly believe that the public health dimension and the public benefit of big data has to be stressed and has to be encouraged."* (Interviewee 12). On a similar note, Interviewee 3 viewed big data as a chance for the scientific community to tackle broad research questions:

> *"I think that the most important benefit is moving away from research on small data packages. I think many questions need a good database and, at the moment... well, research is too widespread, there are too many people who are doing small pieces. I think if you merge these data together you will have a much better chance to have a good research."* (Interviewee 3)

Many interviewees also spoke about the role of big data research in improving prevention and diagnostics. Interviewee 4, however, noted that research participants and patients might not take direct advantage from big data research, but its benefits will be available to the whole of society in the future:

> *"Generally, as I have seen the projects so far, the individual does not benefit directly from the research. Data are used to improve prevention and find new therapy....so the benefit is shifted into the future."* (Interviewee 4)

Likewise, Interviewee 9 commented on the role of big data research to boost precision medicine, in order to find the best treatments for rare diseases and tailor health interventions to specific population sub-groups:

> *"More and more people are moving towards personalized medicine, so having these data at their disposal, they can think of identifying just that small group of people who can have a particular advantage in having a certain therapy."* (Interviewee 9)

When asked about the challenges of big data research and their implications for ethical oversight, the respondents identified a wide range (Figure 6.1). Our interviews, however, revealed a lack of consensus among Cantonal RECs concerning which challenges are the most pressing ("*informed consent*" (Interviewee 13), "*anonymization*" (Interviewee 5), "*results interpretation and generalizability*" (Interviewee 7)). Despite this divergence in prioritizing challenges, the majority of interviewees said that big data research exacerbates privacy and confidentiality risks, potentially resulting in individual and collective harms (*"Huge impact on privacy! Everybody wants you to be under constant surveillance."* (Interviewee 1)). For this reason, respondents stressed the importance of rigorous application of data protection governance and implementation of precautionary measures (e.g., data encryption and anonymization) to secure sensitive information. However, some respondents questioned whether these data protection regulations and practices are actually effective in the context of big data:

> *"...all of a sudden you notice you can find things that you shouldn't have. Big data linkage creates more problems to ensure people' dignity and data confidentiality. These information are precious to people and should not be put in danger."* (Interviewee 13)

> *"Is anonymization possible or not?...Or is it just a word that is not true anymore because it's so easy to identify people behind [the data]?"* (Interviewee 3)

The majority of RECs also mentioned informed consent and research participant autonomy as crucial aspects of ethical research, which are challenged by big data. In fact, REC members argued that informed consent – in its traditional form – is not fit-for-purpose to protect participants in big data research. In particular, Interviewee 8 pointed out that "*the general consent for the use of data for research purposes [which came into force in Switzerland in 2014] is already under review because the first version was not entirely legal and had some flaws.*" On the one hand interviewees seemed to agree that more information should be provided to research participants, especially in light of increased privacy risks. Meaning, it is crucial to "*make transparent to the people what they agree with, which data they disclose, for which purposes data will be used and for how long the data will be stored*" (Interviewee 2). On the other hand, several interviewees questioned the validity of this consent, especially when obtained online in the form of terms of use and e-consent:

> *"It's just hard to agree to a declaration of consent online. These terms and conditions just require you to click and accept, but nobody reads them. That is not an informed and good consent".* (Interviewee 1)

Nonetheless, interviewees' opinions varied concerning which solution could best solve the informed consent impasse. Interviewee 3, for example, commented on the need for a dynamic form of consent. Responsible big data research should allow research participants to choose for which purposes their data are used:

> *"At the moment is all or nothing, we need something much more differentiated. [...] I think to do research in a responsible manner we should say which data are used for what, and give the owner of data or samples the chance to choose each time... I think technically it's possible right now to do it"* (Interviewee 3)

At the opposite side of the spectrum, Interviewee 12, said that projects using biomedical data and providing clear public benefits, should presume participant consent, unless they state otherwise:

> *"The law should be changed to resemble the system they have in Scandinavia where, for research purposes, the access to personal data and samples is guaranteed by a*

presumed consent. Of course, you need to have a democratic and human rights system in place to allow for presumed consent." (Interviewee 12)

Similarly, Interviewee 10 argued from a pragmatic perspective. In the age of big data, it is simply not feasible to obtain informed consent from participants (due to data volume and potential reuses), let alone for big data reuses or retrospectively. Therefore, researchers should focus on obtaining consent only when collecting sensitive data. When using other data types of data (such as data publicly available online), researchers should rely instead on a consent waiver:

"It is just not feasible to obtain consent from everyone. Hundred thousand of people. [...] You just can't. But I am not so sure we need consent for all data. […] We should only protect sensitive data, hence make sure we obtain consent for those. [...] People freely "leave their traces" around the web, giving their information for free to companies while using apps and online services without being concerned. Why should researchers be more concerned?" (Interviewee 10)

Many interviewees explicitly articulated the difficulty to balance the risks and benefits of big data research. They felt particularly uncertain about how to estimate the risks, and justified their concerns with various arguments.

First, the exploratory nature of big data research, as well as the numerous possibilities for data linkage, make it very complex to anticipate the risks (*"another issue is the unforeseeable risk...because as of today we can't tell what we're going to find out about that person through big data analysis*" (Interviewee 8)). Related to this point, Interviewee 3 spoke about the incremental risk of managing incidental findings ("*How are you [researcher] dealing with incidental findings? Is there still a possibility to report them back to the patients or not?*"), due to the large volumes of data which are combined and analyzed. This is especially critical since RECs review research intentions, but do not have control over the outcomes ("*we only see a project on paper at the beginning and then actually implement it. That's kind of out of our hands then*" Interviewee 5).

Second, the use of analytics tools like opaque artificial intelligence algorithms "*that nobody at the end understands*" (Interviewee 3) increases the chance for unclear and incorrect data processes. In turn, these processes can result in "*wrong conclusions"* (Interviewee 7): "*moving away from,*

let's say, a kind of research where you are looking for causality we are diving much more into an area where you just look for correlations, which may be coincidences" (Interviewee 3).

Third, the chances of data hacks and de-identification are hard to anticipate, especially when private companies are involved "*and there are a lot of secrets around these [data protection strategies]. [...] how do you evaluate the quality of protection when you don't know how much they are subject to attacks, how many of those attacks are successful, and what are the steps taken against those attacks?*" (Interviewee 13).

Finally, the presence of private actors in the big data research field makes determining whether the benefits of research are fairly distributed more complex; "*we have to be careful about the fact that big data [research] is not a way of monetizing on our data by big companies...you see it already... they take all of our data and so on and make profit out of that while it is a public good*" (Interviewee 12)

Given the broad spectrum of potential but unclear risks emerging in big data research, the majority of interviewees were dubious when asked to define the threshold of minimal risk:

> *"I mean... if you see the potential for data abuse which is here...and what has already happened...then you can't even speak of minimal risk!"* (Interviewee 1)

Needs of RECs in big data research

Overall, committee members agreed about not having sufficient experience or expertise in technical areas, such as big data analytics or computer science. These weaknesses emerged when trying to understand ("*We can't understand that at all*" (Interviewee 7)) or assess (("*with big data research we simply do not have the know-how*" (Interviewee 6)) biomedical big data projects. Interviewees' concerns predominantly centered around the speed at which new technology evolves. This constant change makes it virtually impossible to have sufficient experience and insight to judge projects with a high degree of certainty, going beyond what seems plausible:

> *"new algorithms with artificial intelligence... I think this is... Well, I have no experience with this and how to deal with this in the future this is an open question."* (Interviewee 3)

Despite this limitation, consensus emerged across RECs regarding their role as key oversight mechanisms for biomedical research, including research relying on big data. While a minority of respondents defended the current way of practicing ethical review and the adequacy of the current laws (*"In terms of principle yes, we are adequately equipped to evaluate biomedical big data projects. We successfully apply them."* (Interviewee 10)), the large majority acknowledged several limitations of the oversight mechanism. When asked about their needs and envisioned solutions, REC members discussed a series of options which could help to address current limitations. Interviewees spoke about solutions at three levels: training, procedures, and regulations.

Regarding the first level, almost all respondents expressed an urgent need to fill the expertise gap within RECs. Expertise was recognized as a crucial factor in effectively fulfilling the oversight mandate. To achieve this goal, interviewees expressed interest in targeted trainings that discuss the characteristics, risks, and ethical implications of big data projects and artificial intelligence applications in biomedicine. In the extract below, Interviewee 4 suggested conducting these trainings in a dynamic format, involving REC members, offering case studies and mock projects to analyze. Meanwhile, other respondents further highlighted that improving REC members' knowledge and allowing for greater exchange could increase review standardization within and across committees:

> *"it would be really good to show examples of how the projects are built and which algorithms are on the back of the analysis and how to they are put together [...] I need concrete examples... [...]The case studies should come from the people who do this...the researchers....to get a proper understanding. [...] then, there should be a discussion among the ethics committees on how to deal with these case studies"* (Interviewee 4)

In addition to these trainings, all respondents confirmed the benefits of involving specialists (in the fields of big data analytics, computer science, and data management) for consultations when needed ("*if I see a problem or so then we get the appropriate expertise. We also do this for quite ordinary applications where the risk cannot be assessed with certainty*" (Interviewee4)). However, the suggestion of Interviewee 6 to include technical experts into RECs was unpopular among the other interviewees:

> "*I think they have to be members, so that we do not have to go and get an expert for an opinion every time*" (Interviewee 6)

> "*I am not convinced that introducing a technical figure can be a solution. Rather get training for the whole committee*." (Interviewee 9)

Concerning solutions at the second level, namely the procedural one, REC members clearly rejected the need for new high-level ethical guidelines ("*Don't draft them. That's my main recommendation*" (Interviewee 12); "*I still can't see exactly how this is supposed to help us*" (Interviewee 4)). In fact, interviewees agreed on having adequate solid research ethics regulation and ethical principles to follow already in place. Rather, several respondents stressed the need for implementable procedures to effectively assess whether big data projects are good or not. On this point Interviewee 5 said that such procedures could "*help to understand how such a big data project should be structured*". Similarly, Interviewee 6 commented:

> *"It is important to define what is good big data research, what must be done when conducting this type of research, and what is optional [...], which methodology is acceptable and which unacceptable".* (Interviewee 6)

Some respondents (like Interviewee 3) emphasized that researchers, too, would benefit from clearer standards about how to ensure data protection, handle unexpected results, certify the validity and quality of the methodology, as well as be transparent about the research question:

> *"I think we have to look whereby certain standards are fulfilled and sometimes we get research application where it's not clear what is the research question... and a lot of tests are done [...] ...and this is then sold as new scientific knowledge... But at the end it may be complete nonsense."* (Interviewee 3)

Interviewee 12 further explained that the REC's attitude toward researchers is not intended to be that of watchdogs who seek to reject research projects simply because they rely on big data. On the contrary, RECs are responsible for promoting ethically aligned research, and want to work together with researchers to improve their projects:

> *"[you need] a lot of goodwill as well to look for solutions. You know... it's not something that we take lightly [reviewing projects]...it's never a matter of saying 'we going to stop this project because uses big data'...The opposite!...We say: 'OK let's look what these researchers want to do and how can we do it in the best way so that they do not hurt people."* (Interviewee 12)

When asked about who should develop these practical guidelines (both those for researchers and for RECs), respondents listed a variety of bodies, starting with Swissethics, followed by the Swiss Academy of Medical Sciences (SAMS) and the Central Ethic Committee (CEC), in collaboration with research institutions and experts (both in ethics and science). If ethics review practices were amended or introduced at an international level, and made valid across countries, REC members would expect the World Health Organization (WHO) to formulate them.

At the third level of solutions (the regulatory), interviewees discussed whether the scope and mandate of RECs should be expanded and also cover those big data projects that currently fall outside their purview. From the respondents' perspectives, these projects might still carry negative consequences for individual health and wellbeing, as well as for broader society. Although REC members had a favorable view of the option of expanding the REC mandate, they also highlighted two crucial points. First, the aim of RECs should align with society's expectations and values. Therefore, the political and health authorities, as well as RECs, should engage with society to define the boundaries of RECs' scope (namely what should be reviewed or not) in light of new technological advancements. Only as a result of this democratic debate, should the law be adjusted – if needed.

> *"Our role is to protect the individual and to decide what is in the interest of a society... Difficult!*
> *I think we need some certain common ground [...] Committees should agree with the society about what should be permitted and what not [...] we need a clear and harmonized understanding of the role of RECs and what is legally required."* (Interviewee 3)

However, although the society may identify a number of core values to respect and promote (e.g., "*privacy, accountability, transparency, public participation"* (Interviewee12)), Interviewee 1 suggested that societal expectations about what exactly ought to be done with the data might remain vague. This is because, "*we live in a pluralistic society today. [...] like here in*

Switzerland...people in Ticino may have different expectations than in Geneva, and the Eastern Switzerland is considered rather 'conservative'." (Interviewee 1).

The second point is of a pragmatic nature. As already mentioned, RECs do not have the capacity and the expertise to review highly technical studies, or studies outside the biomedical field. Therefore, most respondents agreed on the idea of introducing specific oversight boards which could assess the technical features of projects involving big data and artificial intelligence. These boards could complement – rather than substitute for – RECs, and find their place inside research institutions or academia, alongside those already supervising data uses (such as data protection legal offices and data safety monitoring boards).

> *I can imagine that an external body with certain skills could be useful....to evaluate the technical aspects that we do not consider [...] It could be the at the polytechnic...so that with its skills complements our evaluation.* (Interviewee 8)

> *"Possibly on the long term we are going to need something like 'big data board' ...possibly. I do not think they could replace RECs...they will be rather complementary. We could work together to improve and streamline the process."* (Interviewee 10)

Although this idea was endorsed by many, Interviewee 1 called attention to the risk of jeopardizing the efforts of RECs by adding more oversight mechanisms:

> *"This should be carefully considered [...]. I always struggle with too many parallel structures [...]...in the end we have a forest of ethical institutions and nobody knows anymore what is really well reviewed."* (Interviewee 1)

Overall, respondents agree that Cantonal RECs and their current practices have room for improvement, in order to be truly effective and valuable even in times of big data. To succeed in this task however, the good will of RECs alone will not suffice. Rather, the interviewees specified that to successfully tackle the weak points of the current oversight model, Swiss regulators and policymakers should consider these gaps and further clarify the role of RECs among other ethical oversight mechanisms in place.

6.5 Strengths and Limitations

While the methodology of qualitative interview analysis allows for detailed exploration of opinions and perspectives on a given topic, the same study design challenges the generalization of the conclusions. However, although the findings of this study are confined to the Swiss context, the fact that we interviewed members of all seven Cantonal RECs made it possible to represent the full spectrum of perspectives and cultural variation that exist within the country. Furthermore, since the Swiss ethical oversight mechanism partially resembles those existing in other European countries (e.g., in Switzerland's neighboring countries – namely Germany, Austria, France, Italy) and internationally, a certain degree of generalization of results could be justified, upon appropriate consideration.

In this study, a selection bias may have arisen from including only the views of Cantonal RECs. Although other ethics committees exist in Switzerland (e.g., the national ethics committee and the institutional review committees within universities), this study focused on big data research in the biomedical and health field, which is usually reviewed by Cantonal ethics committees. The fact that only the chairperson/vice-chairperson and one scientific secretary per REC were interviewed may also have introduced a bias into the study. Nevertheless, one must consider that chairpersons are those who, in practice, set the agenda for the committee, and the scientific secretaries the ones who first review and evaluate the research protocols. Therefore, we believe that their perspectives and comments have provided valuable insights into the ethics of research with big data in biomedical and health settings.

Finally, the fact that the interviews were conducted before the outbreak of the COVID-19 pandemic can be interpreted as a limitation. Indeed, the recent pandemic has increased pressure on RECs, especially for reviewing public health projects that leverage the power of big data and AI. Nonetheless, the results that this study provides transcend the temporality of current research conditions, as they relate to the complex oversight system of Cantonal RECs, which is not evolving as rapidly. Perhaps, future research could explore whether some processes and functions of Cantonal RECs have changed as a result of the COVID-19 pandemic.

6.6 Discussion

Our findings reveal four main areas of ethical significance. First, the lack of specific normative standards for the ethics review of big data studies. Second, epistemic challenges faced by REC members, specifically insufficient experience and expertise. Third, normative ethical challenges related to the scope of ethical reflection on big data, as several conceptual tools traditionally used to assess biomedical research appeared increasingly inadequate to assess unforeseeable and novel risks generated by big data studies. Finally, proposals for reform emerged from our analysis, including both conservative reforms (e.g. building capacity and promoting data literacy among REC members) and more radical reforms, such as complementing RECs with data-focused oversight bodies. In the following, we provide a detailed analysis of these themes.

Lack of specific review standards

Although REC members share a general idea of what constitutes big data, they lack a precise common definition and clear guidance on how to recognize these studies in practice. This generates uncertainty about whether a research project involves the use of big data. As previous studies indicated, REC members' uncertainty could result in inconsistencies across committees (20). It is relevant to notice, however, that disagreement on a definition of big data is secondary to a lack of tailored standards for reviewing big data research. Our results, in line with previous research, highlight the lack of specific ethical guidelines for evaluating big data projects, and thus the application of traditional ethical frameworks in the evaluation of all projects without distinction (32). Some of our interviewees expressed concern about having to interpret and judge big data research on a case-by-case basis, without guidance that might orient and harmonize their decisions. Furthermore, RECs' diverging interpretations could result in inconsistent evaluations and decisions across committees, which could negatively affect researchers' trust in the oversight system, data sharing practices, and research collaborations (33-35). Although a lack of transparency about evaluation procedures and inconsistencies across RECs' judgments are not exclusive to big data research (36), our findings shows that these limitations in REC practices continue to hamper research.

Limited experience and expertise

REC members acknowledged their limited experience in dealing with big data projects and inadequate expertise about the fundamental technical aspects characterizing these studies. Regarding the first, the REC members recognized that their narrow mandate diminishes their oversight function in big data research. In fact, the narrow boundaries of HRA result in only a portion of big data projects conducted at the national level coming to their attention (37). However, unless the law is amended to expand the purview mission of the ethical oversight mechanism, the Cantonal RECs have no choice but to invite researchers to submit their research voluntarily. In this regard, some studies have suggested that RECs should engage researchers in a dialogue to make them aware of and more accountable for the consequences of their research (38). Concerning the latter – namely the RECs' insufficient expertise to evaluate the technical features of big data research and its challenges – our interviewees seemed well aware of their shortcomings. It should be noted that most REC members expressed both the willingness and commitment to implement strategies to overcome these limitations. Yet, the rapidity with which AI technology and big data applications evolve further complicates the RECs' attempt to get up to speed (39, 40).

Scope of ethical reflection

Third, our findings reveal that REC members are overall well informed about the benefits and challenges brought about by the advent of big data and data analytics techniques. However, they disagreed on which challenges are the most pressing, and more importantly, which tools are best suited to address them. The fact that many interviewees focused on how to adapt and improve the informed consent tool, and implement in the most rigorous way the existing data protection regulation, may signal a problem. In the literature, some authors flagged the risk of viewing these tools as ethics panacea (41, 42). While regulating data re-uses and operationalizing informed consent remain unresolved issues, privacy-focused ethical oversight may be insufficient to address other challenges raised by big data, concerning e.g., justice, dignity, and fairness (43, 44). Our results highlight this gap in the current ethical oversight, as respondents expressed concern about how to balance the risks and benefits of projects. The traditional ethics tools used to assess biomedical research are inadequate and ineffective when trying to assess unforeseeable and novel risks (45). This concern, which remains unresolved for the time being, underscores the need for a broader conversation at the societal level about the importance of big data research, and uses that should be promoted in view of collective interests (46).

Proposals for reform

Our results shed light on the limitations of the current mechanism of the Cantonal RECs, in terms of skills, practices, and guidelines. The REC members – aware of these shortcomings – suggested possible solutions to tackle them. The interviewees' request for training and workshops on big data and AI reveals interest in expanding and boosting their knowledge on relevant topics. In addition, the practice of involving experts to fill RECs' expertise gap can be seen as an attempt to offer better assistance to researchers (47). Interviewees' desire to improve the *status quo* of ethical review is further evidenced by their suggestion of creating complementary oversight mechanisms to the REC. These mechanisms (e.g., big data boards) could review the technical aspects of projects and highlight inherent risks, while keeping pace with the fast-changing nature of research. Some interviewees imagined these boards serving as an accreditation mechanism, to certify the quality of a project's technical features. These boards could operate across disciplines, to certify research conducted both in private and public sectors, regardless of the data types and sources used. Consequently, fewer big data projects would be left without any sort of oversight. Finally, our interviewees strongly defended the role of RECs as a key mechanism for ethical review in research and spoke against overturning the entire ethics review system by introducing new high-level principles or laws. Nevertheless, REC members would welcome more operationalizable guidance on what constitutes a good big data project. Therefore, future research and initiatives should aim to fill this gap by offering ERCs practical guidance for orienting their judgment in the field of big data research.

6.7 Best Practice

Swiss Cantonal RECs should be reformed if they are to strengthen their role and be effective in their practice in the big data research context. In this paper we argue that these reforms should involve not only the practices of REC members, but also their expertise, as well as the regulations that define the mandate of RECs. Other ethics oversight mechanisms outside the Swiss context might benefit from similar revisions as well. In addition, this study suggest that researchers should be proactive in reaching out for RECs' opinions and should be more aware of their responsibilities when conducting research. However, the efforts of researchers must be supported by a system of clear rules and ethics training put in place by a network of actors (such as policymakers, ethics committees, research institutes, universities, and funding bodies) (48).

6.8 Research agenda

In this paper, we reported the perspectives of Swiss cantonal RECs on the challenges they face in reviewing big data projects and their needs in order to adequately address these challenges. We believe this analysis contributes significantly to the existing literature as it is the first qualitative study to survey Swiss RECs about their experiences and views on this topic. Interestingly, our results (collected in the Swiss context) align with the literature produced at the international level. Nevertheless, more research is required to explore the need for globally shared ethical standards for conducting research with big data. In fact, as interdisciplinary and cross-country big data projects increase, the scientific community may need not only clear common data governance, but also a shared vision about what an ethically aligned big data project consists of (49). The recent COVID-19 pandemic exemplified how divergent laws governing research, unclear ethical evaluation methods, and unrobust oversight mechanisms can slow research processes, jeopardize efforts to benefit public health, and reduce public trust in scientific institutions (50).

6.9 Educational Implications

Our results emphasize the need for knowledge exchange and a more productive engagement among the various actors involved in big data research. These include and are not limited to RECs, researchers, research institutions, private enterprises, citizen science groups, and the public (12). In particular, if on the one hand REC members should acquire more technical skills about e.g., data analysis methodologies and AI-enabled technologies, on the other hand researchers should be more informed about the value of and the necessary steps for conducting research ethically. The dynamics of collaboration between RECs and researchers should not only be aimed at fulfilling the requirements imposed by law (i.e., ensuring compliance), but also at increasing mutual knowledge through an open dialogue and positive attitude towards learning. Scholars have argued that positive (although maybe not perfect) actions and responsible big data research can emerge only by asking difficult questions and through transparent confrontation on diverging perspectives (51). Finally, our research findings indicate the crucial importance of informing society about issues related to big data and the use of AI in research. Starting with this democratic engagement, the general public can clarify their expectations regarding research with big data, and thus inform the decisions of other actors involved.

Acknowledgements

We are grateful to our interviewees who kindly spent their time to take part in this research. We also would like to thank Dorothee Caminiti who assisted MI and AF during the completion of the interviews. We are grateful to Shannon Hubbs for her editorial suggestions.

Declaration of interest statement

The authors declare no conflict of interest

References

1. Leonelli S. Scientific Research and Big Data: Edward N. Zalta (ed.); 2020 [updated 2020-05-29. Available from: https://plato.stanford.edu/entries/science-big-data/#BigDataRiskEthiDataScie.
2. Fuller M. Big data and the Facebook scandal: Issues and responses. Theology. 2019;122(1):14-21.
3. Rennie S, Buchbinder M, Juengst E, Brinkley-Rubinstein L, Blue C, Rosen DL. Scraping the Web for Public Health Gains: Ethical Considerations from a 'Big Data'Research Project on HIV and Incarceration. Public Health Ethics. 2020;13(1):111-21.
4. Ferretti A, Ienca M, Hurst S, Vayena E. Big Data, Biomedical Research, and Ethics Review: New Challenges for IRBs. Ethics & human research. 2020;42(5):17-28.
5. Friesen P, Redman B, Caplan A. Of Straws, Camels, Research Regulation, and IRBs. Therapeutic innovation & regulatory science. 2019;53(4):526-34.
6. Vayena E, Gasser U, Wood AB, O'Brien D, Altman M. Elements of a new ethical framework for big data research. Washington and Lee Law Review Online. 2016;72(3).
7. Samuel G, Chubb J, Derrick G. Boundaries Between Research Ethics and Ethical Research Use in Artificial Intelligence Health Research. Journal of Empirical Research on Human Research Ethics. 2021:15562646211002744.
8. Car J, Sheikh A, Wicks P, Williams MS. Beyond the hype of big data and artificial intelligence: building foundations for knowledge and wisdom. BioMed Central; 2019.
9. Blasimme A, Vayena E. The ethics of AI in biomedical research, patient care and public health. Patient Care and Public Health (April 9, 2019) Oxford Handbook of Ethics of Artificial Intelligence, Forthcoming. 2019.
10. Ngan OM, Kelmenson AM. Using Big Data Tools to Analyze Digital Footprint in the COVID-19 Pandemic: Some Public Health Ethics Considerations. Asia Pacific Journal of Public Health. 2020:1010539520984360.
11. Murray CJ, Alamro NMS, Hwang H, Lee U. Digital public health and COVID-19. The Lancet Public Health. 2020;5(9):e469-e70.
12. Vayena E, Gasser U. Strictly biomedical? Sketching the ethics of the big data ecosystem in biomedicine. The ethics of biomedical big data: Springer; 2016. p. 17-39.
13. Price WN, Cohen IG. Privacy in the age of medical big data. Nature medicine. 2019;25(1):37-43.
14. Velarde MCR, Tsantoulis P, Burton-Jeangros C, Aceti MA, Chappuis P, Hurst S. Citizens' Views on Sharing Their Health Data: the Role of Competence, Reliability and Pursuing the Common Good. 2020.
15. Friesen P, Douglas-Jones R, Marks M, Pierce R, Fletcher K, Mishra A, et al. Governing AI-Driven Health Research: Are IRBs Up to the Task? Ethics & Human Research. 2021;43(2):35-42.
16. Metcalf J, Crawford K. Where are human subjects in big data research? The emerging ethics divide. Big Data & Society. 2016;3(1):2053951716650211.
17. Ferretti A, Ienca M, Sheehan M, Blasimme A, Dove ES, Farsides B, et al. Ethics review of big data research: What should stay and what should be reformed? BMC Medical Ethics. 2021;22(1):1-13.
18. Vayena E, Blasimme A. Health research with big data: Time for systemic oversight. The journal of law, medicine & ethics. 2018;46(1):119-29.
19. Parasidis E, Pike E, McGraw D. A belmont report for health data. The New England journal of medicine. 2019;380(16):1493-5.
20. Favaretto M, De Clercq E, Schneble CO, Elger BS. What is your definition of Big Data? Researchers' understanding of the phenomenon of the decade. PloS one. 2020;15(2):e0228987.

21. Nebeker C, Harlow J, Espinoza Giacinto R, Orozco-Linares R, Bloss CS, Weibel N. Ethical and regulatory challenges of research using pervasive sensing and other emerging technologies: IRB perspectives. AJOB empirical bioethics. 2017;8(4):266-76.
22. Hibbin RA, Samuel G, Derrick GE. From "a fair game" to "a form of covert research": Research ethics committee members' differing notions of consent and potential risk to participants within social media research. Journal of Empirical Research on Human Research Ethics. 2018;13(2):149-59.
23. Samuel G, Derrick G, van Leeuwen TN. Ethical challenges around the use of social media data: Views of researchers and research ethics committee members. 2018.
24. Vayena E. Value from health data: European opportunity to catalyse progress in digital health. The Lancet. 2021;397(10275):652-3.
25. Coordination Office for Human Research The Human Research Act and the Ethics Committees for Research 2019 [Available from: https://www.kofam.ch/filemanager/downloads/gutachten_und_berichte/BAG_Faktenblatt_Ethik_A4_E_def.pdf.
26. Swissethics. Mission statement of swissethics 2021 [Available from: https://swissethics.ch/en/.
27. Swiss Federal Office of Public Health (FOPH). Swiss Federal Human Research Act (HRA) 2011 [Available from: https://fedlex.data.admin.ch/filestore/fedlex.data.admin.ch/eli/cc/2013/617/20200101/en/pdf-a/fedlex-data-admin-ch-eli-cc-2013-617-20200101-en-pdf-a.pdf.
28. Favaretto M, De Clercq E, Briel M, Elger BS. Working Through Ethics Review of Big Data Research Projects: An Investigation into the Experiences of Swiss and American Researchers. Journal of Empirical Research on Human Research Ethics. 2020;15(4):339-54.
29. Sims JM. A brief review of the Belmont report. Dimensions of critical care nursing. 2010;29(4):173-4.
30. Emanuel EJ, Wendler D, Killen J, Grady C. What Makes Clinical Research in Developing Countries Ethical? The Benchmarks of Ethical Research. The Journal of Infectious Diseases. 2004;189(5):930-7.
31. Beauchamp TL. The 'four principles' approach to health care ethics. Principles of health care ethics. 2007;29:3-10.
32. Ienca M, Ferretti A, Hurst S, Puhan M, Lovis C, Vayena E. Considerations for ethics review of big data health research: A scoping review. PloS one. 2018;13(10):e0204937.
33. Dove ES, Garattini C. Expert perspectives on ethics review of international data-intensive research: Working towards mutual recognition. Research Ethics. 2018;14(1):1-25.
34. van den Hoonaard WC, Hamilton A. The ethics rupture: Exploring alternatives to formal research-ethics review. 2016.
35. Ballantyne A, Schaefer GO. Public interest in health data research: laying out the conceptual groundwork. Journal of medical ethics. 2020;46(9):610-6.
36. Lynch HF, Abdirisak M, Bogia M, Clapp J. Evaluating the quality of research ethics review and oversight: a systematic analysis of quality assessment instruments. AJOB Empirical Bioethics. 2020:1-15.
37. von Elm E, Briel M. Survey on researchers' opinion about and experience with the Swiss Federal Act on Research involving Human Beings 2019 [Available from: https://www.bag.admin.ch/dam/bag/it/dokumente/biomed/forschung-am-menschen/forschung-biomedizin/Hauptbericht-Befragung.pdf.download.pdf/190715_EDFI_Main%20Report.pdf.
38. Holland K. 18. Enriching Ethics Review Processes in the Spirit of Participatory Dialogue. The Ethics Rupture: University of Toronto Press; 2018. p. 353-75.
39. Prosperi M, Bian J. Is it time to rethink institutional review boards for the era of big data? Nature Machine Intelligence. 2019;1(6):260-.

40. Nebeker C, Torous J, Ellis RJB. Building the case for actionable ethics in digital health research supported by artificial intelligence. BMC medicine. 2019;17(1):1-7.
41. Babb S. Regulating Human Research. Regulating Human Research: Stanford University Press; 2020.
42. Corrigan O. Empty ethics: the problem with informed consent. Sociology of health & illness. 2003;25(7):768-92.
43. Ballantyne A. Adjusting the focus: a public health ethics approach to data research. Bioethics. 2019;33(3):357-66.
44. McKeown A, Mourby M, Harrison P, Walker S, Sheehan M, Singh I. Ethical Issues in Consent for the Reuse of Data in Health Data Platforms. Science and Engineering Ethics. 2021;27(1):1-21.
45. Sheehan M, Thompson R, Fistein J, Davies J, Dunn M, Parker M, et al. Authority and the future of consent in population-level biomedical research. Public health ethics. 2019;12(3):225-36.
46. McMahon A, Buyx A, Prainsack B. Big data governance needs more collective responsibility: the role of harm mitigation in the governance of data use in medicine and beyond. Medical law review. 2020;28(1):155-82.
47. Huh-Yoo J, Kadri R, Buis LR. Pervasive Healthcare IRBs and Ethics Reviews in Research: Going Beyond the Paperwork. IEEE Pervasive Computing. 2021;20(1):40-4.
48. Samuel G, Derrick GE, van Leeuwen T. The ethics ecosystem: Personal ethics, network governance and regulating actors governing the use of social media research data. Minerva. 2019;57(3):317-43.
49. ÓhÉigeartaigh SS, Whittlestone J, Liu Y, Zeng Y, Liu Z. Overcoming barriers to cross-cultural cooperation in AI ethics and governance. Philosophy & Technology. 2020;33(4):571-93.
50. Gardner L, Ratcliff J, Dong E, Katz A. A need for open public data standards and sharing in light of COVID-19. The Lancet Infectious Diseases. 2020.
51. Zook M, Barocas S, Boyd D, Crawford K, Keller E, Gangadharan SP, et al. Ten simple rules for responsible big data research. Public Library of Science San Francisco, CA USA; 2017.

CHAPTER 7: Ethics Review of Big Data Research: What Should Stay and What Should Be Reformed?

Published as: Ferretti, A., Ienca, M., Sheehan, M., Blasimme, A., Dove, E. S., Farsides, B., Friesen, P., Kahn, J., Karlen, W., Kleist, P., Liao, S.M., Nebeker, C., Samuel, G., Shabani, M., Rivas Velarde, M., Vayena, E. Ethics review of big data research: what should stay and what should be reformed? (2021) *BMC Medical Ethics,* doi: 10.1186/s12910-021-00616-4

7.1 Abstract

Ethics review is the process of assessing the ethics of research involving humans. The Ethics Review Committee (ERC) is the key oversight mechanism designated to ensure ethics review. Whether or not this governance mechanism is still fit for purpose in the data-driven research context remains a debated issue among research ethics experts.

In this article, we seek to address this issue in a twofold manner. First, we review the strengths and weaknesses of ERCs in ensuring ethical oversight. Second, we map these strengths and weaknesses onto specific challenges raised by big data research. We distinguish two categories of potential weakness. The first category concerns persistent weaknesses, i.e., those which are not specific to big data research, but may be exacerbated by it. The second category concerns novel weaknesses, i.e., those which are created by and inherent to big data projects. Within this second category, we further distinguish between purview weaknesses related to the ERC's scope (e.g., how big data projects may evade ERC review) and functional weaknesses, related to the ERC's way of operating. Based on this analysis, we propose reforms aimed at improving the oversight capacity of ERCs in the era of big data science.

We believe the oversight mechanism could benefit from these reforms because they will help to overcome data-intensive research challenges and consequently benefit research at large.

7.2 Background

The debate about the adequacy of the Ethics Review Committee (ERC) as the chief oversight body for big data studies is partly rooted in the historical evolution of the ERC. Particularly relevant is the ERC's changing response to new methods and technologies in scientific research. ERCs – also known as Institutional Review Boards (IRBs) or Research Ethics Committees (RECs) – came to existence in the 1950s and 1960s (1). Their original mission was to protect the interests of human research participants, particularly through an assessment of potential harms to them (e.g., physical pain or psychological distress) and benefits that might accrue from the proposed research. ERCs expanded in scope during the 1970s, from participant protection towards ensuring valuable and ethical human subject research (e.g., having researchers implement an informed consent process), as well as supporting researchers in exploring their queries (2).

Fast forward fifty years, and a lot has changed. Today, biomedical projects leverage unconventional data sources (e.g., social media), partially inscrutable data analytics tools (e.g., machine learning), and unprecedented volumes of data (3-5). Moreover, the evolution of research practices and new methodologies such as post-hoc data mining have blurred the concept of '*human subject'* and elicited a shift towards the concept of *data subject* - as attested in data protection regulations. (6, 7). With data protection and privacy concerns being in the spotlight of big data research review, language from data protection laws has worked its way into the vocabulary of research ethics. This terminological shift further reveals that big data, together with modern analytic methods used to interpret the data, creates novel dynamics between researchers and participants (8). Research data repositories about individuals and aggregates of individuals are considerably expanding in size. Researchers can remotely access and use large volumes of potentially sensitive data without communicating or actively engaging with study participants. Consequently, participants become more vulnerable and subjected to the research itself (9). As such, the *nature of risk* involved in this new form of research changes too. In particular, it moves from the risk of physical or psychological harm towards the risk of informational harm, such as *privacy breaches* or *algorithmic discrimination* (10). This is the case, for instance, with projects using data collected through web search engines, mobile and smart devices, entertainment websites, and social media platforms. The fact that health-related research is leaving hospital labs and spreading into online space creates novel opportunities for research, but also raises novel challenges for ERCs. For this reason, it is important to re-examine the fit between new data-driven forms of research and existing oversight mechanisms (11).

The suitability of ERCs in the context of big data research is not merely a theoretical puzzle but also a practical concern resulting from recent developments in data science. In 2014, for example, the so-called 'emotional contagion study' received severe criticism for avoiding ethical oversight by an ERC, failing to obtain research consent, violating privacy, inflicting emotional harm, discriminating against data subjects, and placing vulnerable participants (e.g., children and adolescents) at risk (12, 13). In both public and expert opinion (14), a responsible ERC would have rejected this study because it contravened the research ethics principles of preventing harm (in this case, emotional distress) and adequately informing data subjects. However, the protocol adopted by the researchers was not required to undergo ethics review under US law (15) for two reasons. First, the data analyzed were considered non-identifiable, and researchers did not engage directly with subjects, exempting the study from ethics review. Second, the study team included both scientists affiliated with a public university (Cornell) and Facebook employees. The

affiliation of the researchers is relevant because—in the US and some other countries—privately funded studies are not subject to the same research protections and ethical regulations as publicly funded research (16). An additional example is the 2015 case in which the United Kingdom (UK) National Health Service (NHS) shared 1.6 million pieces of identifiable and sensitive data with Google DeepMind. This data transfer from the public to the private party took place legally, without the need for patient consent or ethics review oversight (17). These cases demonstrate how researchers can pursue potentially risky big data studies without falling under the ERC's purview. The limitations of the regulatory framework for research oversight are evident, in both private and public contexts.

The gaps in the ERC's regulatory process, together with the increased sophistication of research contexts—which now include a variety of actors such as universities, corporations, funding agencies, public institutes, and citizens associations—has led to an increase in the range of oversight bodies. For instance, besides traditional university ethics committees and national oversight committees, funding agencies and national research initiatives have increasingly created internal ethics review boards (18, 19). New participatory models of governance have emerged, largely due to an increase in subjects' requests to control their own data (20). Corporations are creating research ethics committees as well, modelled after the institutional ERC (21). In May 2020, for example, Facebook welcomed the first members of its Oversight Board, whose aim is to review the company's decisions about content moderation (22). Whether this increase in oversight models is motivated by the urge to fill the existing regulatory gaps, or whether it is just 'ethics washing', is still an open question. However, other types of specialized committees have already found their place alongside ERCs, when research involves international collaboration and data sharing (23). Among others, data safety monitoring boards, data access committees, and responsible research and innovation panels serve the purpose of covering research areas left largely unregulated by current oversight (24).

The data-driven digital transformation challenges the purview and efficacy of ERCs. It also raises fundamental questions concerning the role and scope of ERCs as the oversight body for ethical and methodological soundness in scientific research[1]. Among these questions, this article will

[1] There is an unsettled discussion about whether ERCs ought to play a role in evaluating both scientific and ethical aspects of research, or whether these can even come apart - but we will not go into detail here. 25. Dawson AJ, Yentis SM. Contesting the science/ethics distinction in the review of clinical research. Journal of Medical Ethics. 2007;33(3):165-7, 26. Angell EL, Bryman A, Ashcroft RE, Dixon-Woods M. An analysis of decision letters by research ethics committees: the ethics/scientific quality boundary examined. BMJ Quality & Safety. 2008;17(2):131-6.

explore whether ERCs are still capable of their intended purpose, given the range of novel (maybe not categorically new, but at least different in practice) issues that have emerged in this type of research. To answer this question, we explore some of the challenges that the ERC oversight approach faces in the context of big data research and review the main strengths and weaknesses of this oversight mechanism. Based on this analysis, we will outline possible solutions to address current weaknesses and improve ethics review in the era of big data science.

7.3 Main text

Strengths of the ethics review via ERC

Historically, ERCs have enabled cross disciplinary exchange and assessment (27). ERC members typically come from different backgrounds and bring their perspectives to the debate; when multi-disciplinarity is achieved, the mixture of expertise provides the conditions for a solid assessment of advantages and risks associated with new research. Committees which include members from a variety of backgrounds are also suited to promote projects from a range of fields, and research that cuts across disciplines (28). Within these committees, the reviewers' expertise can be paired with a specific type of content to be reviewed. This one-to-one match can bring timely and, ideally, useful feedback (29). In many countries (e.g., European countries, the United States (US), Canada, Australia), ERCs are explicitly mandated by law to review many forms of research involving human participants; moreover, these laws also describe how such a body should be structured and the purview of its review (30, 31). In principle, ERCs also aim to be representative of society and the research enterprise, including members of the public and minorities, as well as researchers and experts (32). And in performing a gatekeeping function to the research enterprise, ERCs play an important role: they recognize that both experts and lay people should have a say, with different views to contribute (33).

Furthermore, the ERC model strives to ensure independent assessment. The fact that ERCs assess projects "from the outside" and maintain a certain degree of objectivity towards what they are reviewing, reduces the risk of overlooking research issues and decreases the risk for conflicts of interest. Moreover, being institutionally distinct—for example, being established by an organization that is distinct from the researcher or the research sponsor—brings added value to the research itself as this lessens the risk for conflict of interest. Conflict of interest is a serious issue in research ethics because it can compromise the judgment of reviewers. Institutionalized

review committees might particularly suffer from political interference. This is the case, for example, for universities and health care systems (like the NHS), which tend to engage "in house" experts as ethics boards members. However, ERCs that can prove themselves independent are considered more trustworthy by the general public and data subjects; it is reassuring to know that an independent committee is overseeing research projects (34).

The ex-ante (or pre-emptive) ethical evaluation of research studies is by many considered the standard procedural approach of ERCs (35). Though the literature is divided on the usefulness and added value provided by this form of review (36, 37), ex-ante review is commonly used as a mechanism to ensure the ethical validity of a study design before the research is conducted (38, 39). Early research scrutiny aims at risk-mitigation: the ERC evaluates potential research risks and benefits, in order to protect participants' physical and psychological well-being, dignity, and data privacy. This practice saves researchers' resources and valuable time by preventing the pursuit of unethical or illegal paths (40). Finally, the ex-ante ethical assessment gives researchers an opportunity to receive feedback from ERCs, whose competence and experience may improve the research quality and increase public trust in the research (41).

All *strengths* mentioned in this section are strengths of the ERC model in principle. In practice, there are many ERCs that are not appropriately interdisciplinary or representative of the population and minorities, that lack independence from the research being reviewed, and that fail to improve research quality, and may in fact hinder it. We now turn to consider some of these weaknesses in more detail.

Weaknesses of the ethics review via ERC

In order to assess whether ERCs are adequately equipped to oversee big data research, we must consider the weaknesses of this model. We identify two categories of weaknesses which are described in the following section and summarized in Figure 7.1:

- *Persistent weaknesses*: those existing in the current oversight system, which could be exacerbated by big data research
- *Novel weaknesses*: those brought about by and specific to the nature of big data projects

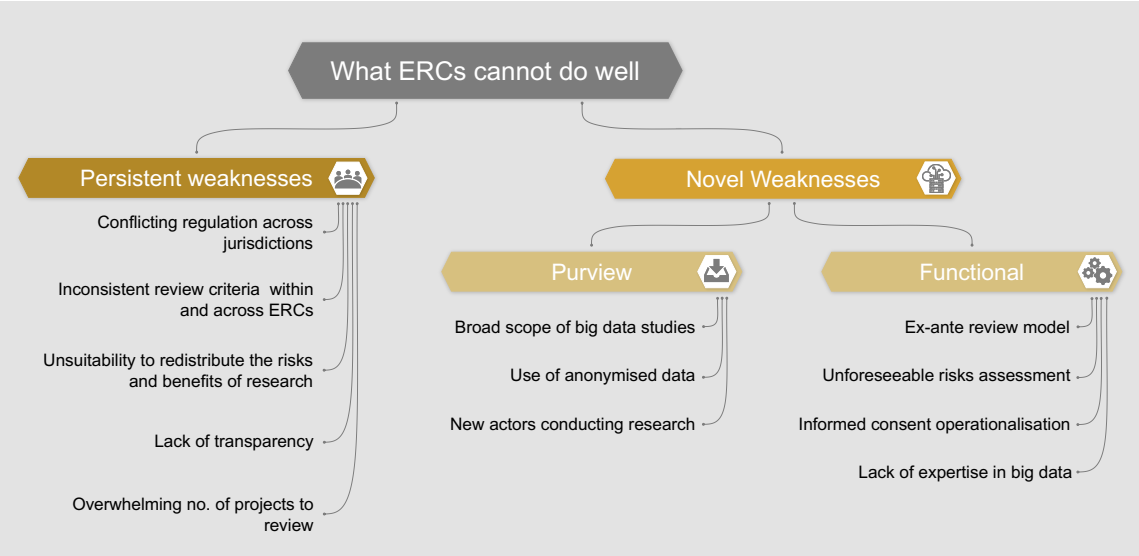
Within this second category of novel weaknesses, we further differentiate between:

- *Purview weaknesses*: reasons why some big data projects may bypass the ERCs' purview

- *Functional weaknesses*: reasons why some ERCs may be inadequate to assess big data projects specifically

We base the conceptual distinction between persistent and novel weaknesses on the fact that big data research diverges from traditional biomedical research in many respects. As previously mentioned, big data projects are often broad in scope, involve new actors, use unprecedented methodologies to analyze data, and require specific expertise. Furthermore, the peculiarities of big data itself (e.g., being large in volume and from a variety of sources) make data-driven research different in practice from traditional research. However, we should not consider the category of "novel weaknesses" a closed category. We do not argue that weaknesses mentioned here do not, at least partially, overlap with others which already exist. In fact, in almost all cases of 'novelty', (i) there is some link back to a concept from traditional research ethics, and (ii) some thought has been given to the issue outside of a big data or biomedical context (e.g., the problem of ERCs' expertise has arisen in other fields (42)). We believe that by creating conceptual clarity about novel oversight challenges presented by big data research, we can begin to identify tailored reforms.

Figure 7.1: Weaknesses of the ERCs



Persistent weaknesses

As regulation for research oversight varies between countries, ERCs often suffer from a lack of harmonization. This weakness in the current oversight mechanism is compounded by big data research, which often relies on multi-center international consortia. These consortia in turn depend on approval by multiple oversight bodies demanding different types of scrutiny (43). Furthermore, big data research may give rise to collaborations between public bodies, universities, corporations, foundations, and citizen science cooperatives. In this network, each stakeholder has different priorities and depends upon its own rules for regulation of the research process (44-46). Indeed, this expansion of regulatory bodies and aims does not come with a coordinated effort towards agreed-upon review protocols (47). The lack of harmonization is perpetuated by academic journals and funding bodies with diverging views on the ethics of big data. If the review bodies which constitute the "ethics ecosystem" (19) do not agree to the same ethics review requirements, a big data project deemed acceptable by an ERC in one country may be rejected by another ERC, within or beyond the national borders.

In addition, there is inconsistency in the assessment criteria used within and across committees. Researchers report subjective bias in the evaluation methodology of ERCs, as well as variations in ERC judgements which are not based on morally relevant contextual considerations (48, 49). Some authors have argued that the probability of research acceptance among experts increases if some research peer or same-field expert sits on the evaluation committee (50, 51). The judgement of an ERC can also be influenced by the boundaries of the scientific knowledge of its members. These boundaries can impact the ERC's approach towards risk taking in unexplored fields of research (52). Big data research might worsen this problem since the field is relatively new, with no standardized metric to assess risk within and across countries (53). The committees do not necessarily communicate with each other to clarify their specific role in the review process, or try to streamline their approach to the assessment. This results in unclear oversight mandates and inconsistent ethical evaluations (27, 54).

Additionally, ERCs may fall short in their efforts to justly redistribute the risks and benefits of research. The current review system is still primarily tilted toward protecting the interests of individual research participants. ERCs do not consistently assess societal benefit, or risks and benefits in light of the overall conduct of research (balancing risks for the individual with collective benefits). Although demands on ERCs vary from country to country (55), the ERC approach is still generally tailored towards traditional forms of biomedical research, such as clinical trials and

longitudinal cohort studies with hospital patients. These studies are usually narrow in scope and carry specific risks only for the participants involved. In contrast, big data projects can impact society more broadly. As an example, computational technologies have shown potential to determine individuals' sexual orientation by screening facial images (56). An inadequate assessment of the common good resulting from this type of study can be socially detrimental (57). In this sense, big data projects resemble public health research studies, with an ethical focus on the common good over individual autonomy (58). Within this context, ERCs have an even greater responsibility to ensure the just distribution of research benefits across the population. Accurately determining the social value of big data research is challenging, as negative consequences may be difficult to detect before research begins. Nevertheless, this task remains a crucial objective of research oversight.

The literature reports examples of the failure of ERCs to be accountable and transparent (59). This might be the result of an already unclear role of ERCs. Indeed, the ERCs practices are an outcome of different levels of legal, ethical, and professional regulations, which largely vary across jurisdictions. Therefore, some ERCs might function as peer counselors, others as independent advisors, and still others as legal controllers. What seems to be common across countries, though, is that ERCs rarely disclose their procedures, policies, and decision-making process. The ERCs' "secrecy" can result in an absence of trust in the ethical oversight model (60).This is problematic because ERCs rely on public acceptance as accountable and trustworthy entities (61). In big data research, as the number of data subjects is exponentially greater, a lack of accountability and an opaque deliberative process on the part of ERCs might bring even more significant public backlash. Ensuring truthfulness of the stated benefits and risks of research is a major determinant of trust in both science and research oversight. Researchers are another category of stakeholders negatively impacted by poor communication and publicity on the part of the ERC. Commentators have shown that ERCs often do not clearly provide guidance about the ethical standards applied in the research review (62). For instance, if researchers provide unrealistic expectations of privacy and security to data subjects, ERCs have an institutional responsibility to flag those promises (e.g., about data security and the secondary-uses of subject data), especially when the research involves personal and high sensitivity data (63). For their part, however, ERCs should make their expectations and decision-making processes clear.

Finally, ERCs face the increasing issue of being overwhelmed by the number of studies to review (64, 65). Whereas ERCs originally reviewed only human subjects research happening in natural

sciences and medicine, over time they also became the ethical body of reference for those conducting human research in the social sciences (e.g., in behavioral psychology, educational sciences, etc.). This increase in demand creates pressure on ERC members, who often review research *pro bono* and on a voluntary basis. The wide range of big data research could exacerbate this existing issue. Having more research to assess and less time to accomplish the task may negatively impact the quality of the ERC's output, as well as increase the time needed for review (66). Consequently, researchers might carry out potentially risky studies because the relevant ethical issues of those studies were overlooked. Furthermore, research itself could be significantly delayed, until it loses its timely scientific value.

Novel weaknesses: Purview weaknesses

To determine whether the ERC is still the most fit-for-purpose entity to oversee big data research, it is important to establish under which conditions big data projects fall under the purview of ERCs. Historically, research oversight has primarily focused on human subject research in the biomedical field, using public funding. In the US for instance, each review board is responsible for a subtype of research based on content or methodology (for example there are IRBs dedicated to validating clinical trial protocols, assessing cancer treatments, examining pediatric research, and reviewing qualitative research). This traditional ethics review structure cannot accommodate big data research (2). Big data projects often reach beyond a single institution, cut across disciplines, involve data collected from a variety of sources, re-use data not originally collected for research purposes, combine diverse methodologies, orient towards population-level research, rely on large data aggregates, and emerge from collaboration with the private sector. Given this scenario, big data projects may likely fall beyond the purview of ERCs.

Another case in which big data research does not fall under ERC purview is when it relies on anonymized data. If researchers use data that cannot be traced back to subjects (anonymized or non-personal data), then according to both the US Common Rule and HIPAA regulations, the project is considered safe enough to be granted an ethics review waiver. If instead researchers use pseudonymized (or de-identified) data, they must apply for research ethics review, as in principle the key that links the de-identified data with subjects could be revealed or hacked, causing harm to subjects. In the European Union, it would be left to each Member State (and national laws or policies at local institutions) to define whether research using anonymized data should seek ethical review. This case shows once more that current research ethics regulation is relatively loose and disjointed across jurisdictions, and may leave areas where big data research

is unregulated. In particular, the special treatment given anonymized data comes from an emphasis on risk at the individual level. So far in the big data discourse, the concept of harm has been mainly linked to vulnerability in data protection. Therefore if privacy laws are respected, and protection is built into the data system, researchers can prevent harmful outcomes (40). However, this view is myopic as it does not include other misuses of data aggregates, such as group discrimination and dignitary harm. These types of harm are already emerging in the big data ecosystem, where anonymized data reveal health patterns of a certain sub-group, or computational technologies include strong racial biases (67, 68). Furthermore, studies using anonymized data should not be deemed oversight-free by default, as it is increasingly hard to anonymize data. Technological advancements might soon make it possible to re-identify individuals from aggregate data sets (69).

The risks associated with big data projects also increase due to the variety of actors involved in research alongside university researchers (e.g., private companies, citizen science associations, bio-citizen groups, community workers cooperatives, foundations, and non-profit organizations) (70, 71). The novel aspect of health-related big data research compared with traditional research is that anyone who can access large amounts of data about individuals and build predictive models based on that data, can now determine and infer the health status of a person without directly engaging with that person in a research program (72). Facebook, for example, is carrying out a suicide prediction and prevention project, which relies exclusively on the information that users post on the social network (18). Because this type of research is now possible, and the available ethics review model exempts many big data projects from ERC appraisal, gaps in oversight are growing (17, 73). Just as corporations can re-use publicly available datasets (such as social media data) to determine life insurance premiums (74), citizen science projects can be conducted without seeking research oversight (75). Indeed, participant-led big data research (despite being increasingly common) is another area where the traditional overview model is not effective (76). In addition, ERCs might consider research conducted outside academia or publicly funded institutions to be not serious. Thus ERCs may disregard review requests from actors outside the academic environment (e.g., by the citizen science or health tech start up) (77).

Novel weaknesses: Functional weaknesses

Functional weaknesses are those related to the skills, composition, and operational activities of ERCs in relation to big data research.

From this functional perspective, we argue that the ex-ante review model might not be appropriate for big data research. Project assessment at the project design phase or at the data collection level is insufficient to address emerging challenges that characterize big data projects – especially as data, over time, could become useful for other purposes, and therefore be re-used or shared (78). Limitations of the ex-ante review model have already become apparent in the field of genetic research (79). In this context, biobanks must often undergo a second ethics assessment to authorize the specific research use on exome sequencing of their primary data samples (80). Similarly, in a case in which an ERC approved the original collection of sensitive personal data, a data access committee would ensure that the secondary uses are in line with original consent and ethics approval. However, if researchers collect data from publicly accessible platforms, they can potentially use and re-use data for research lawfully, without seeking data subject consent or ERC review. This is often the case in social media research. Social media data, which are collected by researchers or private companies using a form of broad consent, can be re-used by researchers to conduct additional analysis without ERC approval. It is not only the re-use of data that poses unforeseeable risks. The ex-ante approach might not be suitable to assess other stages of the data lifecycle (81), such as deployment machine learning algorithms.

Rather than re-using data, some big data studies build models on existing data (using data mining and machine learning methods), creating new data, which is then used to further feed the algorithms (82). Sometimes it is not possible to anticipate which analytic models or tools (e.g., artificial intelligence) will be leveraged in the research. And even then, the nature of computational technologies which extract meaning from big data make it difficult to anticipate all the correlations that will emerge from the analysis (37). This is an additional reason that big data research often has a tentative approach to a research question, instead of growing from a specific research hypothesis (83).The difficulty of clearly framing the big data research itself makes it even harder for ERCs to anticipate unforeseeable risks and potential societal consequences. Given the existing regulations and the intrinsic exploratory nature of big data projects, the mandate of ERCs does not appear well placed to guarantee research oversight. It seems even less so if we consider problems that might arise after the publication of big data studies, such as repurposing or dual-use issues (84).

ERCs also face the challenge of assessing the value of informed consent for big data projects. To re-obtain consent from research subjects is impractical, particularly when using consumer generated data (e.g., social media data) for research purposes. In these cases, researchers often

rely on broad consent and consent waivers. This leaves the data subjects unaware of their participation in specific studies, and therefore makes them incapable of engaging with the research progress. Therefore, the data subjects and the communities they represent become vulnerable towards potential negative research outcomes. The tool of consent has limitations in big data research—it cannot disclose all possible future uses of data, in part because these uses may be unknown at the time of data generation. Moreover, researchers can access existing datasets multiple times and reuse the same data with alternative purposes (85). What should be the ERCs' strategy, given the current model of informed consent leaves an ethical gap in big data projects? ERCs may be tempted to focus on the consent challenge, neglecting other pressing big data issues (78). However, the literature reports an increasing number of authors who are against the idea of a new consent form for big data studies (5).

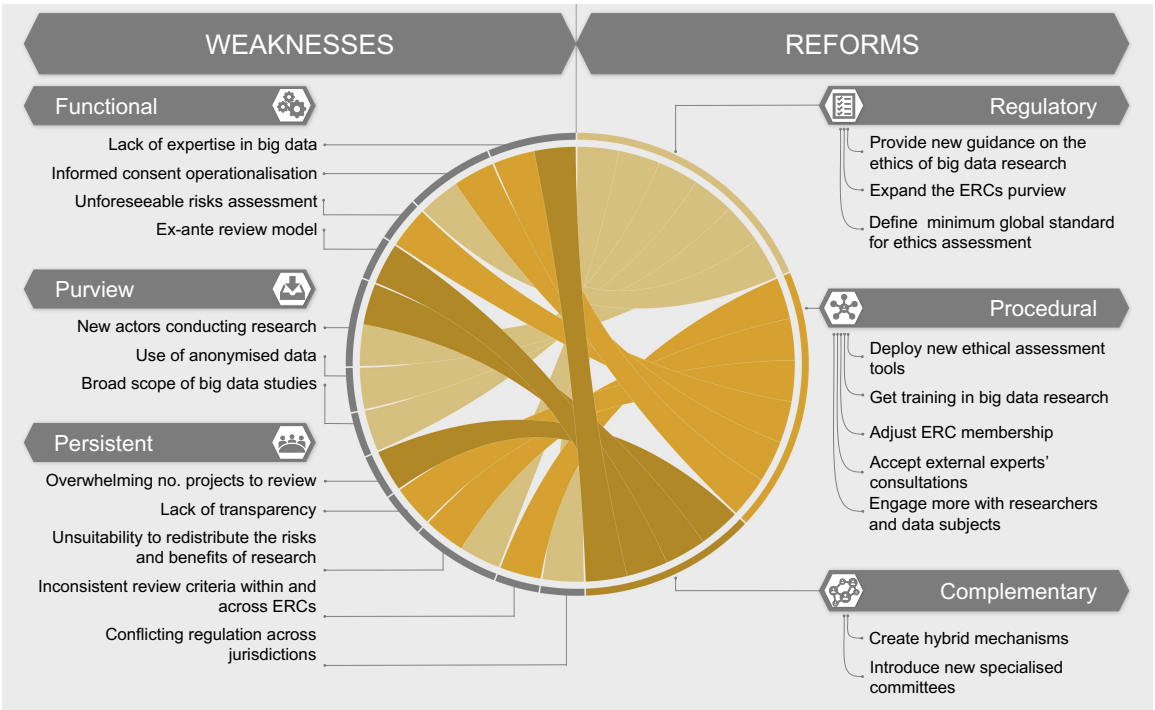
A final widely discussed concern is the ERC's inadequate expertise in the area of big data research (86, 87). In the past, there have been questions about the technical and statistical expertise of ERC members. For example, ERCs have attempted to conform social science research to the clinical trial model, using the same knowledge and approach to review both types of research (88). However, big data research poses further challenges to ERCs' expertise. First, the distinct methodology of big data studies (based on data aggregation and mining) requires a specialized technical expertise (e.g., information systems, self-learning algorithms, and anonymization protocols). Indeed, big data projects have a strong technical component, due to data volume and sources, which brings specific challenges (e.g., collecting data outside traditional protocols on social media) (89, 90). Second, ERCs may be unfamiliar with new actors involved in big data research, such as citizen science actors or private corporations. Because of this lack of relevant expertise, ERCs may require unjustified amendments to research studies, or even reject big data projects *tout-court* (36). Finally, ERCs may lose credibility as an oversight body capable of assessing ethical violations and research misconduct. In the past, ERCs solved this challenge by consulting independent experts in a relevant field when reviewing a protocol in that domain. However, this solution is not always practical as it depends upon the availability of an expert. Furthermore, experts may be researchers working and publishing in the field themselves. This scenario would be problematic because researchers would have to define the rules experts must abide by, compromising the concept of independent review (19). Nonetheless, this problem does not disqualify the idea of expertise but requires high transparency standards regarding rule development and compliance. Other options include *ad-hoc* expert committees or provision of

relevant training for existing committee members (47, 91, 92). Given these options, which one is best to address ERCs' lack of expertise in big data research?

Reforming the ERC

Our analysis shows that ERCs play a critical role in ensuring ethical oversight and risk-benefit evaluation (93), assessing the scientific validity of a project in its early stages, and offering an independent, critical, and interdisciplinary approach to the review. These strengths demonstrate why the ERC is an oversight model worth holding on to. Nevertheless, ERCs carry persistent big data-specific weaknesses, reducing their effectiveness and appropriateness as oversight bodies for data-driven research. To answer our initial research question, we propose that the current oversight mechanism is not as fit for purpose to assess the ethics of big data research as it could be in principle. ERCs should be improved at several levels to be able to adequately address and overcome these challenges. Changes could be introduced at the level of the regulatory framework as well as procedures. Additionally, reforming the ERC model might mean introducing complementary forms of oversight. In this section we explore these possibilities. Figure 7.2 offers an overview of the reforms that could aid ERCs in improving their process.

Figure 7.2: Reforms overview for the research oversight mechanism



Regulatory reforms

The regulatory design of research oversight is the first aspect which needs reform. ERCs could benefit from new guidance (e.g., in the form of a flowchart) on the ethics of big data research. This guidance could build upon a deep rethinking of the importance of data for the functioning of societies, the way we use data in society, and our justifications for this use. In the UK, for instance, individuals can generally opt out of having their data (e.g., hospital visit data, health records, prescription drugs) stored by physicians' offices or by NHS digital services. However, exceptions to this opt-out policy apply when uses of the data are vital to the functioning of society (for example, in the case of official national statistics or overriding public interest, such as the COVID-19 pandemic) (94).

We imagine this new guidance also re-defining the scope of ERC review, from protection of individual interest to a broader research impact assessment. In other words, it will allow the ERC's scope to expand and to address purview issues which were previously discussed. For example, less research will be oversight-free because more factors would trigger ERC purview in the first place. The new governance would impose ERC review for research involving anonymized data, or big data research within public-private partnerships. Furthermore, ERC purview could be extended beyond the initial phase of the study to other points in the data lifecycle (95). A possible option is to assess a study after its conclusion (as is the case in the pharmaceutical industry): ERCs could then decide if research findings and results should be released and further used by the scientific community. This new ethical guidance would serve ERCs not only in deciding whether a project requires review, but also in learning from past examples and best practices how to best proceed in the assessment. Hence, this guidance could come in handy to increase transparency surrounding assessment criteria used across ERCs. Transparency could be achieved by defining a minimum global standard for ethics assessment that allows international collaboration based on open data and a homogenous evaluation model. Acceptance of a global standard would also mean that the same oversight procedures will apply to research projects with similar risks and research paths, regardless of whether they are carried on by public or private entities. Increased clarification and transparency might also streamline the review process within and across committees, rendering the entire system more efficient.

Procedural reforms

Procedural reforms might target specific aspects of the ERC model to make it more suitable for the review of big data research. To begin with, ERCs should develop new operational tools to mitigate emerging big data challenges. For example, the AI Now algorithmic impact assessment tool, which appraises the ethics of automated decision systems, and informs decisions about whether or not to deploy the systems in society, could be used (96). Forms of broad consent (97) and dynamic consent (20) can also address some of the issues raised, by using, re-using, and sharing big data (publicly available or not). Nonetheless, informed consent should not be considered a panacea for all ethical issues in big data research – especially in the case of publicly available social media data (98). If the ethical implications of big data studies affect the society and its vulnerable sub-groups, individual consent cannot be relied upon as an effective safeguard. For this reason, ERCs should move towards a more democratic process of review. Possible strategies include engaging research subjects and communities in the decision-making process or promoting a co-governance system. The recent Montreal Declaration for Responsible AI is an example of an ethical oversight process developed out of public involvement (99). Furthermore, this inclusive approach could increase the trustworthiness of the ethics review mechanism itself (100). In practice, the more that ERCs involve potential data subjects in a transparent conversation about the risks of big data research, the more socially accountable the oversight mechanism will become.

ERCs must also address their lack of big data and general computing expertise. There are several potential ways to bridge this gap. First, ERCs could build capacity with formal training on big data. ERCs are willing to learn from researchers about social media data and computational methodologies used for data mining and analysis (86). Second, ERCs could adjust membership to include specific experts from needed fields (e.g., computer scientists, biotechnologists, bioinformaticians, data protection experts). Third, ERCs could engage with external experts for specific consultations. Despite some resistance to accepting help, recent empirical research has shown that ERCs may be inclined to rely upon external experts in case of need (87).

In the data-driven research context, ERCs must embrace their role as regulatory stewards, and walk researchers through the process of ethics review (40). ERCs should establish an open communication channel with researchers to communicate the value of research ethics while clarifying the criteria used to assess research. If ERCs and researchers agree to mutually increase transparency, they create an opportunity to learn from past mistakes and prevent future

ones (101). Universities might seek to educate researchers on ethical issues that can arise when conducting data-driven research. In general, researchers would benefit from training on identifying issues of ethics or completing ethics self-assessment forms, particularly if they are responsible for submitting projects for review (102). As biomedical research is trending away from hospitals and clinical trials, and towards people's homes and private corporations, researchers should strive towards greater clarity, transparency, and responsibility. Researchers should disclose both envisioned risks and benefits, as well as the anticipated impact at the individual and population level (54). ERCs can then more effectively assess the impact of big data research and determine whether the common good is guaranteed. Furthermore, they might examine how research benefits are distributed throughout society. Localized decision making can play a role here (55). ERCs may take into account characteristics specific to the social context, to evaluate whether or not the research respects societal values.

Complementary reforms

An additional measure to tackle the novelty of big data research might consist in reforming the current research ethics system through regulatory and procedural tools. However, this strategy may not be sufficient: the current system might require additional support from other forms of oversight to complement its work.

One possibility is the creation of hybrid review mechanisms and norms, merging valuable aspects of the traditional ERC review model with more innovative models, which have been adopted by various partners involved in the research (e.g., corporations, participants, communities) (103). This integrated mechanism of oversight would cover all stages of big data research and involve all relevant stakeholders (104). Journals and the publishing industry could play a role within this hybrid ecosystem in limiting potential dual use concerns. For instance, in the research publication phase, resources could be assigned to editors so as to assess research integrity standards and promote only those projects which are ethically aligned. However, these implementations can have an impact only when there is a shared understanding of best practice within the oversight ecosystem (19).

A further option is to include specialized and distinct ethical committees alongside ERCs, whose purpose is to assess big data research and provide sectorial accreditation to researchers. In this model, ERCs would not be overwhelmed by the numbers of study proposals to review and could outsource evaluations requiring specialist knowledge in the field of big data. It is true that

specialized committees (data safety monitoring boards, data access committees, and responsible research and innovation panels) already exist and support big data researchers in ensuring data protection (e.g., system security, data storage, data transfer). However, something like a “data review board” could assess research implications both for the individual and society, while reviewing a project’s technical features. Peer review could play a critical role in this model: the research community retains the expertise needed to conduct ethical research and to support each other when the path is unclear (102).

Despite their promise, these scenarios all suffer from at least one primary limitation. The former might face a backlash when attempting to bring together the priorities and ethical values of various stakeholders, within common research norms. Furthermore, while decentralized oversight approaches might bring creativity over how to tackle hard problems, they may also be very dispersive and inefficient. The latter could suffer from overlapping scope across committees, resulting in confusing procedures, and multiplying efforts while diluting liability. For example, research oversight committees have multiplied within the United States, leading to redundancy and disharmony across committees (47). Moreover, specialized big data ethics committees working in parallel with current ERCs could lead to questions over the role of the traditional ERC, when an increasing number of studies will be big data studies.

7.4 Conclusions

ERCs face several challenges in the context of big data research. In this article, we sought to bring clarity regarding those which might affect the ERC’s practice, distinguishing between novel and persistent weaknesses which are compounded by big data research. While these flaws are profound and inherent in the current sociotechnical transformation, we argue that the current oversight model is still partially capable of guaranteeing the ethical assessment of research. However, we also advance the notion that introducing reform at several levels of the oversight mechanism could benefit and improve the ERC system itself. Among these reforms, we identify the urgency for new ethical guidelines and new ethical assessment tools to safeguard society from novel risks brought by big data research. Moreover, we recommend that ERCs adapt their membership to include necessary expertise for addressing the research needs of the future. Additionally, ERCs should accept external experts’ consultations and consider training in big data technical features as well as big data ethics. A further reform concerns the need for transparent

engagement among stakeholders. Therefore, we recommend that ERCs involve both researchers and data subjects in the assessment of big data research. Finally, we acknowledge the existing space for a coordinated and complementary support action from other forms of oversight. However, the actors involved must share a common understanding of best practice and assessment criteria in order to efficiently complement the existing oversight mechanism. We believe that these adaptive suggestions could render the ERC mechanism sufficiently agile and well-equipped to overcome data-intensive research challenges and benefit research at large.

Abbreviations

ERC(s): Ethics Review Committee(s)
HIPAA: Health Insurance Portability and Accountability Act
IRB(s): Institutional Review Board(s)
NHS: National Health Service
REC(s): Research Ethics Committee(s)
UK: United Kingdom
US: United States

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

Not applicable

Competing interests

The authors declare that they have no competing interests

Funding

This research is supported by the Swiss National Science Foundation under award 407540_167223 (NRP 75 Big Data). MS1 is grateful for funding from the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre (BRC). The funding bodies did not take part in designing this research and writing the manuscript.

Authors' contributions

AF drafted the manuscript, MI, MS1 and EV contributed substantially to the writing. EV is the senior lead on the project from which this article derives. All the authors (AF, MI, MS1, AB, ESD,

BF, PF, JK, WK, PK, SML, CN, GS, MS2, MRV, EV) contributed greatly to the intellectual content of this article, edited it, and approved the final version.

Acknowledgements

This article reports the ideas and the conclusions emerged during a collaborative and participatory online workshop. All authors participated in the “Big Data Challenges for Ethics Review Committees” workshop, held online the 23-24 April 2020 and organized by the Health Ethics and Policy Lab, ETH Zurich.

References

1. Moon MR. The History and Role of Institutional Review Boards: A Useful Tension. AMA Journal of Ethics. 2009;11(4):311-6.
2. Friesen P, Kearns L, Redman B, Caplan AL. Rethinking the Belmont report? The American Journal of Bioethics. 2017;17(7):15-21.
3. Nebeker C, Torous J, Ellis RJB. Building the case for actionable ethics in digital health research supported by artificial intelligence. BMC medicine. 2019;17(1):137.
4. Ienca M, Ferretti A, Hurst S, Puhan M, Lovis C, Vayena E. Considerations for ethics review of big data health research: A scoping review. PloS one. 2018;13(10).
5. Hibbin RA, Samuel G, Derrick GE. From “a fair game” to “a form of covert research”: Research ethics committee members’ differing notions of consent and potential risk to participants within social media research. Journal of empirical research on human research ethics. 2018;13(2):149-59.
6. Maldoff G. How GDPR changes the rules for research: International Association of Privacy Protection; 2020 [Available from: https://iapp.org/news/a/how-gdpr-changes-the-rules-for-research/.
7. Samuel G, Buchanan E. Guest Editorial: Ethical Issues in Social Media Research. SAGE Publications Sage CA: Los Angeles, CA; 2020. p. 3-11.
8. Shmueli G. Research Dilemmas with Behavioral Big Data. Big Data. 2017;5(2).
9. Sula CA. Research ethics in an age of big data. Bulletin of the association for information science and technology. 2016;42(2):17-21.
10. Metcalf J, Crawford K. Where are human subjects in Big Data research? The emerging ethics divide. Big Data & Society. 2016;3(1):2053951716650211.
11. Vayena E, Gasser U, Wood AB, O'Brien D, Altman M. Elements of a new ethical framework for big data research. Washington and Lee Law Review Online. 2016;72(3).
12. Goel V. As Data Overflows Online, Researchers Grapple With Ethics: The New York Times; 2014 [Available from: https://www.nytimes.com/2014/08/13/technology/the-boon-of-online-data-puts-social-science-in-a-quandary.html.
13. Vitak J, Shilton K, Ashktorab Z, editors. Beyond the Belmont principles: Ethical challenges, practices, and beliefs in the online data research community. Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing; 2016.
14. BBC World. Facebook emotion experiment sparks criticism 2014 [Available from: https://www.bbc.com/news/technology-28051930.
15. Fiske ST, Hauser RM. Protecting human research participants in the age of big data. Proceedings of the National Academy of Sciences of the United States of America. 2014;111(38):13675.
16. Klitzman R, Appelbaum PS. Facebook’s emotion experiment: Implications for research ethics: The Hastings Center; 2014 [Available from: https://www.thehastingscenter.org/facebooks-emotion-experiment-implications-for-research-ethics/.
17. Ballantyne A, Stewart C. Big Data and Public-Private Partnerships in Healthcare and Research. Asian Bioethics Review. 2019;11(3):315-26.
18. Barnett I, Torous J. Ethics, transparency, and public health at the intersection of innovation and Facebook's suicide prevention efforts. American College of Physicians; 2019.
19. Samuel G, Derrick GE, van Leeuwen T. The ethics ecosystem: Personal ethics, network governance and regulating actors governing the use of social media research data. Minerva. 2019;57(3):317-43.

20. Vayena E, Blasimme A. Biomedical big data: new models of control over access, use and governance. Journal of bioethical inquiry. 2017;14(4):501-13.
21. BBC World. Google announces AI ethics panel: BBC World; 2019 [Available from: https://www.bbc.com/news/technology-47714921.
22. Clegg N. Welcoming the Oversight Board - About Facebook: FACEBOOK; 2020 [updated 2020-05-06. Available from: https://about.fb.com/news/2020/05/welcoming-the-oversight-board/.
23. Shabani M, Dove ES, Murtagh M, Knoppers BM, Borry P. Oversight of genomic data sharing: what roles for ethics and data access committees? Biopreservation and biobanking. 2017;15(5):469-74.
24. Joly Y, Dove ES, Knoppers BM, Bobrow M, Chalmers D. Data sharing in the post-genomic world: the experience of the International Cancer Genome Consortium (ICGC) Data Access Compliance Office (DACO). PLoS Comput Biol. 2012;8(7):e1002549.
25. Dawson AJ, Yentis SM. Contesting the science/ethics distinction in the review of clinical research. Journal of Medical Ethics. 2007;33(3):165-7.
26. Angell EL, Bryman A, Ashcroft RE, Dixon-Woods M. An analysis of decision letters by research ethics committees: the ethics/scientific quality boundary examined. BMJ Quality & Safety. 2008;17(2):131-6.
27. Nichols AS. Research ethics committees (RECS)/institutional review boards (IRBS) and the globalization of clinical research: Can ethical oversight of human subjects research be standardized. Wash U Global Stud L Rev. 2016;15:351.
28. Garrard E, Dawson A. What is the role of the research ethics committee? Paternalism, inducements, and harm in research ethics. Journal of medical ethics. 2005;31(7):419-23.
29. Page SA, Nyeboer J. Improving the process of research ethics review. Research Integrity and Peer Review. 2017;2(1):14.
30. Bowen AJ. Models of institutional review board function. 2008.
31. McGuinness S. Research ethics committees: the role of ethics in a regulatory authority. Journal of Medical Ethics. 2008;34(9):695-700.
32. Kane C, Takechi K, Chuma M, Nokihara H, Takagai T, Yanagawa H. Perspectives of non-specialists on the potential to serve as ethics committee members. Journal of International Medical Research. 2019;47(5):1868-76.
33. Kirkbride J, George A. Lay REC members: patient and public. Journal of Medical Ethics. 2020;39(12):780-2.
34. Resnik DB. Trust as a Foundation for Research with Human Subjects. The Ethics of Research with Human Subjects: Protecting People, Advancing Science, Promoting Trust. Cham: Springer International Publishing; 2018. p. 87-111.
35. Kritikos M. Research Ethics Governance: The European Situation. Handbook of Research Ethics and Scientific Integrity. 2020:33-50.
36. Molina JL, Borgatti SP. Moral bureaucracies and social network research. Social Networks [Internet]. 2019 16.11.2020.
37. Sheehan M, Dunn M, Sahan K. Reasonable disagreement and the justification of pre-emptive ethics governance in social research: a response to Hammersley. Journal of Medical Ethics. 2018;44:719-20.
38. Mustajoki H. Pre-emptive research ethics: Finnish NationalBoard on Research Integrity Tenk; 2018 [Available from: https://vastuullinentiede.fi/en/doing-research/pre-emptive-research-ethics.
39. Biagetti M, Gedutis A. Towards Ethical Principles of Research Evaluation in SSH. The Third Research Evaluation in SSH Conference, Valencia, 19-20 September 20192019. p. 19-20.

40. Dove ES. Regulatory Stewardship of Health Research: Edward Elgar Publishing; 2020.
41. Tene O, Polonetsky J. Beyond IRBs: Ethical guidelines for data research. Washington and Lee Law Review Online. 2016;72(3):458.
42. Bloss C, Nebeker C, Bietz M, Bae D, Bigby B, Devereaux M, et al. Reimagining human research protections for 21st century science. Journal of medical Internet research. 2016;18(12):e329.
43. Dove ES, Garattini C. Expert perspectives on ethics review of international data-intensive research: Working towards mutual recognition. Research Ethics. 2018;14(1):1-25.
44. van den Broek T, van Veenstra AF. Governance of big data collaborations: How to balance regulatory compliance and disruptive innovation. Technological Forecasting and Social Change. 2018;129:330-8.
45. Jackman M, Kanerva L. Evolving the IRB: building robust review for industry research. Washington and Lee Law Review Online. 2016;72(3):442.
46. Someh I, Davern M, Breidbach CF, Shanks G. Ethical issues in big data analytics: A stakeholder perspective. Communications of the Association for Information Systems. 2019;44(1):34.
47. Friesen P, Redman B, Caplan A. Of Straws, Camels, Research Regulation, and IRBs. Therapeutic innovation & regulatory science. 2019;53(4):526-34.
48. Kohn T, Shore C. The ethics of university ethics committees. Risk management and the research imagination, in Death of the public university. 2017:229-49.
49. Friesen P, Yusof ANM, Sheehan M. Should the Decisions of Institutional Review Boards Be Consistent? Ethics & human research. 2019;41(4):2-14.
50. Binik A, Hey SP. A framework for assessing scientific merit in ethical review of clinical research. Ethics & human research. 2019;41(2):2-13.
51. Derrick GE, Haynes A, Chapman S, Hall WD. The association between four citation metrics and peer rankings of research influence of Australian researchers in six fields of public health. PLoS One. 2011;6(4):e18521.
52. Luukkonen T. Conservatism and risk-taking in peer review: Emerging ERC practices. Research Evaluation. 2012;21(1):48-60.
53. Dove ES, Townend D, Meslin EM, Bobrow M, Littler K, Nicol D, et al. Ethics review for international data-intensive research. Science. 2016;351(6280):1399-400.
54. Abbott L, Grady C. A systematic review of the empirical literature evaluating IRBs: What we know and what we still need to learn. Journal of Empirical Research on Human Research Ethics. 2011;6(1):3-19.
55. Shaw DM, Elger BS. The relevance of relevance in research. Swiss Medical Weekly. 2013;143(1920).
56. Kosinski Y, Wang M. Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. Journal of Personality and Social Psychology. 2018;114(2):246-57.
57. Levin S. LGBT groups denounce 'dangerous' AI that uses your face to guess sexuality: The Guardian; 2017 [updated 2017-09-09. Available from: http://www.theguardian.com/world/2017/sep/08/ai-gay-gaydar-algorithm-facial-recognition-criticism-stanford.
58. Tan S, Zhao Y, Huang W. Neighborhood Social Disadvantage and Bicycling Behavior: A Big Data-Spatial Approach Based on Social Indicators. Social Indicators Research. 2019;145(3):985-99.
59. Lynch HF. Opening closed doors: Promoting IRB transparency. The Journal of Law, Medicine & Ethics. 2018;46(1):145-58.

60. Samuel GN, Farsides B. Public trust and 'ethics review'as a commodity: the case of Genomics England Limited and the UK's 100,000 genomes project. Medicine, Health Care and Philosophy. 2018;21(2):159-68.
61. Nebeker C, Lagare T, Takemoto M, Lewars B, Crist K, Bloss CS, et al. Engaging research participants to inform the ethical conduct of mobile imaging, pervasive sensing, and location tracking research. Translational behavioral medicine. 2016;6(4):577-86.
62. Clapp JT, Gleason KA, Joffe S. Justification and authority in institutional review board decision letters. Social Science & Medicine. 2017;194:25-33.
63. Sheehan M, Friesen P, Balmer A, Cheeks C, Davidson S, Devereux J, et al. Trust, trustworthiness and sharing patient data for research. Journal of Medical Ethics [Internet]. 2020.
64. Klitzman R. The ethics police?: The struggle to make human research safe: Oxford University Press; 2015.
65. Cantonal Ethics Committee Zurich. Annual Report 2019. 2019 [Available from: https://www.zh.ch/content/dam/zhweb/bilder-dokumente/organisation/gesundheitsdirektion/ethikkommission-/jahresberichte-kek/Jahresbericht_KEK%20ZH%202019_09-03-2020_PKL.pdf.
66. Lynch HF, Abdirisak M, Bogia M, Clapp J. Evaluating the quality of research ethics review and oversight: a systematic analysis of quality assessment instruments. AJOB Empirical Bioethics. 2020:1-15.
67. Hoffman S. What genetic testing teaches about long-term predictive health analytics regulation. 2019.
68. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science. 2019;366(6464):447-53.
69. Yoshiura H. Re-identifying people from anonymous histories of their activities. 2019 IEEE 10th International Conference on Awareness Science and Technology (iCAST); 23-25 Oct. 20192019. p. 1-5.
70. Holm S, Ploug T. Big Data and Health Research—The Governance Challenges in a Mixed Data Economy. Journal of Bioethical Inquiry. 2017;14(4):515-25.
71. Nebeker C. mHealth Research Applied to Regulated and Unregulated Behavioral Health Sciences. The Journal of Law, Medicine & Ethics. 2020;48(1_suppl):49-59.
72. Marks M. Emergent Medical Data: Health Information Inferred by Artificial Intelligence. UC Irvine Law Review (2021, Forthcoming). 2020.
73. Friesen P, Douglas Jones R, Marks M, Pierce R, Fletcher K, Mishra A, et al. Governing AI-driven health research: are IRBs up to the task? Ethics & Human Research. 2020 Forthcoming
74. Baron J. Life Insurers Can Use Social Media Posts To Determine Premiums, As Long As They Don't Discriminate: Forbes; 2019 [Available from: https://www.forbes.com/sites/jessicabaron/2019/02/04/life-insurers-can-use-social-media-posts-to-determine-premiums/.
75. Wiggins A, Wilbanks J. The rise of citizen science in health and biomedical research. The American Journal of Bioethics. 2019;19(8):3-14.
76. Ienca M, Vayena E. "Hunting Down My Son's Killer": New Roles of Patients in Treatment Discovery and Ethical Uncertainty. Journal of Bioethical Inquiry. 2020:1-11.
77. Grant AD, Wolf GI, Nebeker C. Approaches to governance of participant-led research: a qualitative case study. BMJ open. 2019;9(4):e025633.
78. Dove ES, Townend D, Meslin EM, Bobrow M, Littler K, Nicol D, et al. Ethics review for international data-intensive research. Science. 2016;351(6280):1399-400.

79. Mascalzoni D, Hicks A, Pramstaller P, Wjst M. Informed consent in the genomics era. PLoS Med. 2008;5(9):e192.
80. McGuire AL, Beskow LM. Informed consent in genomics and genetic research. Annual review of genomics and human genetics. 2010;11:361-81.
81. Roth S, Luczak-Roesch M. Deconstructing the data life-cycle in digital humanitarianism. Information, Communication & Society. 2020;23(4):555-71.
82. Gal A, Senderovich A. Process Minding: Closing the Big Data Gap. International Conference on Business Process Management: Springer; 2020. p. 3-16.
83. Ferretti A, Ienca M, Hurst S, Vayena E. Big Data, Biomedical Research, and Ethics Review: New Challenges for IRBs. Ethics & human research. 2020;42(5):17-28.
84. Ienca M, Vayena E. Dual use in the 21st century: emerging risks and global governance. Swiss Medical Weekly. 2018;148:w14688.
85. Shabani M, Borry P. Rules for processing genetic data for research purposes in view of the new EU General Data Protection Regulation. European Journal of Human Genetics. 2018;26(2):149-56.
86. Nebeker C, Harlow J, Espinoza Giacinto R, Orozco-Linares R, Bloss CS, Weibel N. Ethical and regulatory challenges of research using pervasive sensing and other emerging technologies: IRB perspectives. AJOB empirical bioethics. 2017;8(4):266-76.
87. Sellers C, Samuel G, Derrick G. Reasoning “uncharted territory”: notions of expertise within ethics review panels assessing research use of social media. Journal of Empirical Research on Human Research Ethics. 2020;15(1-2):28-39.
88. Schrag ZM. The case against ethics review in the social sciences. Research Ethics. 2011;7(4):120-31.
89. Beskow LM, Hammack-Aviran CM, Brelsford KM, O'Rourke PP. Expert Perspectives on Oversight for Unregulated mHealth Research: Empirical Data and Commentary. The Journal of Law, Medicine & Ethics. 2020;48(1_suppl):138-46.
90. Huh-Yoo J, Rader E. It's the Wild, Wild West: Lessons Learned From IRB Members' Risk Perceptions Toward Digital Research Data. Proceedings of the ACM on Human-Computer Interaction. 2020;4(CSCW1):1-22.
91. Research; NHA. Gene Therapy Advisory Committee 2020 [Available from: https://www.hra.nhs.uk/about-us/committees-and-services/res-and-recs/gene-therapy-advisory-committee/.
92. Research; NHA. The Social Care Research Ethics Committee (REC) 2020 [Available from: https://www.hra.nhs.uk/planning-and-improving-research/policies-standards-legislation/social-care-research/.
93. Sheehan M, Dunn M, Sahan K. In defence of governance: ethics review and social research. Journal of medical ethics. 2017;44(10):710-6.
94. NHS UK. When your choice does not apply. 2019 [Available from: https://www.nhs.uk/your-nhs-data-matters/where-your-choice-does-not-apply/.
95. Master Z, Martinson BC, Resnik DB. Expanding the scope of research ethics consultation services in safeguarding research integrity: Moving beyond the ethics of human subjects research. The American Journal of Bioethics. 2018;18(1):55-7.
96. Reisman D, Schultz J, Crawford K, Whittaker M. Algorithmic impact assessments: A practical framework for public agency accountability. AI Now Institute. 2018:1-22.
97. Sheehan M. Broad consent is informed consent. Bmj. 2011;343:d6900.
98. Sheehan M, Thompson R, Fistein J, Davies J, Dunn M, Parker M, et al. Authority and the Future of Consent in Population-Level Biomedical Research. Public Health Ethics. 2019;12(3):225-36.

99. Montréal; Ud. Montréal Declaration for a Responsible Development of Artificial Intelligence 2019 [Available from: https://www.montrealdeclaration-responsibleai.com.
100. McCoy MS, Jongsma KR, Friesen P, Dunn M, Neuhaus CP, Rand L, et al. National Standards for Public Involvement in Research: missing the forest for the trees. Journal of Medical Ethics. 2018;44(12):801-4.
101. Brown C, Spiro J, Quinton S. The role of research ethics committees: Friend or foe in educational research? An exploratory study. British Educational Research Journal. 2020;46(4):747-69.
102. Pagoto S, Nebeker C. How scientists can take the lead in establishing ethical practices for social media research. Journal of the American Medical Informatics Association. 2019;26(4):311-3.
103. Harlow J, Weibel N, Al Kotob R, Chan V, Bloss C, Linares-Orozco R, et al. Using participatory design to inform the Connected and Open Research Ethics (CORE) commons. Science and engineering ethics. 2020;26(1):183-203.
104. Vayena E, Blasimme A. Health research with big data: Time for systemic oversight. The journal of law, medicine & ethics. 2018;46(1):119-29.

SECTION II (B)

ETHICAL OVERSIGHT IN HEALTH APPS

CHAPTER 8: From Principles to Practice: Benchmarking Government Guidance on Health Apps

Peer reviewed and published as: Ferretti, A., Ronchi, E., Vayena, E. (2019). From Principles to Practice: Benchmarking Government Guidance on Health Apps. *The Lancet Digital Health*, doi:10.1016/S2589-7500(19)30027-5.

8.1 Main text

Patient-facing mobile health applications (apps) hold the promise to change the way individuals take responsibility for their own health by enabling more effective delivery of health information, allowing better monitoring of symptoms, and encouraging healthier lifestyles (1). Enthusiasm about the potential of health apps has grown rapidly, generating uncertainty as to who should regulate such apps and how. In most countries, medical device regulation applies only to a subset of high-risk health apps that have well defined medical purposes. However, most health apps available on the market target a wide range of health-related issues, including diet and exercise, pregnancy, and mental health, while still being considered non-medical devices. These apps can collect a variety of personal data (2) and, because they are designed to affect health, it is important to ensure their safety, validity, reliability, privacy, and security. Many low-quality health apps exist that, as well as providing advice that is incomplete, misleading, or wrong, might also fall short of meeting the expected standards in privacy and security (3)(4). In a 2017 report, the Organisation for Economic Co-Operation and Development (OECD) concluded that the use of low-quality, non-medical health apps raises a wide range of ethical, legal, and governance issues, and pointed to the need for international agreement on minimum standards in quality assurance controls (5).

We examined guidance for the development of safe, secure, and reliable apps issued by data protection authorities in nine OECD countries (see Appendix 1 Chapter 8 for methodological details). All these authorities had reported guidance relating to mobile apps in response to the 2017 Census of the International Conference on Data Protection and Privacy Commissioners. We also examined guidance by national health authorities in the same nine countries, and international guidance issued by WHO and the European Commission (EC). We did a comparative assessment of the guidance against the qualitative indicators and the principles and best practices set out in the 2013 OECD Privacy Guidelines and the 2016 OECD Recommendation on Health Data Governance (Appendix 2 Chapter 8).

The documents issued by data protection authorities are largely similar to one another. Most address app developers, providing information in question format or checklists, and are easily accessible online. All the documents from data protection authorities focus mainly on privacy and data protection, and include the core principles of the two OECD recommendations. These

documents aim to clarify how to comply with relevant national legislation for app development in general, but not specifically for non-medical health apps.

Among the documents from data protection authorities, the extent to which explanatory and operational guidance is included varies. Although most documents cover rules on user data (e.g., guidance on user's consent, right to access their data, and data portability), only four documents address privacy impact assessment. The documents also included few specifications concerning third party access and use of data. In particular, the documents did not contain advice on good practice to adopt when data is used for marketing and in-app advertising. Furthermore, few documents discuss good practice on apps for children or disabled people.

We also searched the websites of the national health authorities in the same nine countries for specific guidance, and contacted their respective OECD representatives for confirmation (Appendix 3 Chapter 8). Most European ministries and the Australian department of health confirmed that their countries had not developed specific guidance. Guidance might have been produced at a territory level or by professional organisations, but reviewing such documents was not in the scope of this report. Only two national health authorities have issued specific guidance: the UK National Health Service (NHS) and the French Haute Autorité de Santé (HAS). The types of guidance delivered by these authorities differ substantially. HAS good practice guidelines are aimed at app developers and evaluators (evaluating bodies, consumer associations, or medical professional organisations), and the guidance is based on five categories of good practice requirements: informing users, health content, technical content, security and reliability, and usability. The guidance includes a risk matrix to help tailor and assess good practice according to the app's intended uses. In the UK, the NHS has developed a review process for the selection of non-medical apps to be included in the NHS app library. The review process is based on an online questionnaire listing good practice goals in eight core areas: clinical effectiveness, regulatory approval, clinical safety, data protection, security, usability and accessibility, interoperability, and technical stability.

At the EU level, in 2016, the EC issued the draft Privacy Code of Conduct on health apps for public consultation. This document includes most of the criteria listed by the UK and French health authorities, in addition to specific and practical guidance on privacy and data protection principles to be taken into account by health app developers. At international level, the mHealth Evidence Reporting Assessment, developed by the WHO Technical Evidence Review Group in 2016, also

includes most of the criteria listed by the NHS and HAS, with a focus on assessment of an app's technical features, rather than non-technical aspects such as clinical safety and privacy.

In summary, although governments are developing regulation and guidance for app developers, our report shows that this guidance is siloed and not comprehensive. Data protection authorities are focused on privacy and data protection issues, whereas health authorities introduce safety and efficacy considerations. The spread of different sources of guidance across agencies leaves it up to developers to navigate the complex regulatory environment. As apps are often global products, multiple guidelines and different agencies and requirements make compliance onerous and accountability measures unclear. Although our analysis cannot directly link this fragmented landscape with the poor quality standards of non-medical health apps that has been documented, fragmentation and lack of comprehensive guidance is probably not conducive to effective governance (6).

Professional organisations, academics, and the private sector have stepped in to provide additional guidance. For example, guidelines have been produced by Xcertia to support consumers and clinicians in choosing mobile health apps, and to help developers in complying with industry-wide accepted standards. The British Publicly Available Specification 277 (health and wellness apps—quality criteria across the life cycle) provides a wide variety of specifications on health apps, including topics such as fitness for purpose, risk management, quality criteria, and support. Additionally, medical associations, such as the Royal Dutch Medical Association (7), have provided sector-specific guidance and standards to evaluate effectiveness and safety of health apps. At the national level in the USA, the Agency for Healthcare Research and Quality issued an evaluation of the efficacy, usability, and features of commercially available apps for diabetes self-management in 2018 (8). In the UK, the NHS, in collaboration with the National Institute for Health and Care Excellence, has issued an evidence standards framework. Finally, although standardised voluntary self-certification or star rating tools could help consumers and clinicians to reach informed decisions regarding app use, these tools rely on developers' accountability and competence and have yet to be proven reliable (9).

Greater policy coordination for the governance of health apps is needed to reduce guidance gaps (e.g., on data access and user autonomy), to make quality standards visible and clear, and to create an accessible common reference for developers, users, and payors. As countries move to develop strategies for greater patient-centred care, we must enable cross-country agreement on

minimum quality assurance standards to guide app development and use. An OECD-led multi-stakeholder initiative, building on OECD's ongoing work on health data governance, would be a crucial step towards global consensus.

Acknowledgements

We declare no competing interests. AF is supported by the Swiss National Science Foundation (grant number 407540_167223; NRP 75 Big Data). ER is a Senior Policy Analyst for the Organisation for Economic Co-Operation and Development (OECD). The opinions expressed and arguments employed herein do not necessarily reflect the official views of the OECD or of the governments of its member countries.

References

1. Boulos MNK, Brewer AC, Karimkhani C, Buller DB, Dellavalle RP. Mobile medical and health apps: state of the art, concerns, regulatory control and certification. Online J Public Health Inform. 2014; 5: 229
2. Vayena E, Dzenowagis J, Brownstein JS, Sheikh A. Policy implications of big data in the health sector. Bull World Health Organ. 2018; 96: 66-68
3. Huckvale K, Prieto JT, Tilney M, Benghozi PJ, Car J. Unaddressed privacy risks in accredited health and wellness apps: a cross-sectional systematic assessment. BMC Med. 2015; 13: 214
4. Powell AC, Torous J, Chan S, et al. Interrater reliability of mHealth app rating measures: analysis of top depression and smoking cessation apps. JMIR Mhealth Uhealth. 2016; 4: e15
5. Organisation for Economic Co-operation and Development. Mobile technology-based services for global health and wellness: opportunities and challenges. 2017. Available at: http://www.oecd.org/sti/ieconomy/Summary-Mobile-Health-Oct2016.pdf. Date accessed: April 10, 2019
6. Bates DW, Landman A, Levine DM. Health apps and health policy: what is needed?. JAMA. 2018; 320: 1975-1976
7. Royal Dutch Medical Association. Medical app checker: evaluation of mobile medical apps. 2016. Available at: https://www.knmg.nl/web/file?uuid=39ce54ca-ceec-4d91-b242-05bb8cee5e23&owner=5c945405-d6ca-4deb-aa16-7af2088aa173&contentid=3915. Date accessed: April 10, 2019
8. Agency for Healthcare Research and Quality. Mobile health applications for self-management of diabetes. 2018. Available at: https://effectivehealthcare.ahrq.gov/sites/default/files/pdf/technical-brief-31_mobile-health-applications-for-self-management-of-diabetes.pdf. Date accessed: April 10, 2019
9. Chan SR, Misra S. Certification of mobile apps for health care. JAMA. 2014; 312: 1155-1156

CHAPTER 9: Digital Contact Tracing Against COVID-19 in Europe: Current Features and Ongoing Developments

Submitted for review as: Blasimme, A.*, Ferretti, A.*, Vayena, E. Digital Contact Tracing Against COVID-19 in Europe: Current Features and Ongoing Developments.

* These authors contributed equally and share first authorship

9.1 Abstract

The SARS-CoV-2 pandemic is a public health challenge of unprecedented scale. In the midst of the first wave of the pandemic, governments worldwide introduced digital contact tracing systems as part of a strategy to contain the spread of the virus. In Europe, after intense discussion about privacy-related risks involving policymakers, technology experts, information technology companies, and – albeit to a limited extent – the public at large, technical protocols were created to support the development of privacy-compatible proximity tracing apps. However, as the second wave of SARS-CoV-2 sweeps the continent, digital contact tracing in Europe is evolving in terms of both technological and governance features. To enable policymakers to harness the full potential of digital health tools against SARS-CoV-2, this paper examines the evolution of digital contact tracing in eight European countries. Our study highlights that while privacy and data protection are at the core of contact tracing apps in Europe, countries differ in their technical protocols, and in their capacity to utilize collected data beyond proximity tracing alone. In particular, the most recently released apps tend to offer users more granular information about risk in specific locations, and to collect data about user whereabouts, in order to enhance retrospective contact tracing capacity. These developments signal a shift from a strict interpretation of data minimization and purpose limitation towards a more expansive approach to digital contact tracing in Europe, calling for careful scrutiny and appropriate oversight.

9.2 Introduction

The SARS-CoV-2 pandemic is a public health challenge of unprecedented scale. Worldwide, 100.4 million people have tested positive for SARS-CoV-2 and 2.16 million have lost their life to the coronavirus disease (COVID-19) (1). Europe alone has had almost 18 million cases and 425 thousand deaths, according to the most recent estimates of the European Centre for Disease Control and Prevention (2). Since late summer 2020, Europe has faced a resurgence of new cases as a second wave of SARS-CoV-2 spread across the continent, placing health systems under severe pressure and forcing governments to reinstate restrictions similar to those adopted in the first quarter of the year.

Alongside restrictions to population movement during the first wave of the pandemic, governments throughout the world introduced digital contact tracing (DCT) systems, in the hope that this new digital health technology would help contain the spread of the virus (3). Asian countries were among the first to adopt DCT. Recognizing the public health potential of DCT, many European countries followed suit during the spring, developing national DCT systems in an attempt to expand their contact tracing capability.

Despite the promising potential of DCT, its introduction gave rise to intense debate over ethical, legal, and societal implications (ELSI). In particular, some characteristics of the Asian approach (mandatory use, centralized protocols, GPS- or cell tower-based geolocation) are seen by many as incompatible with European legal provisions and ethical views about the value of individual privacy.

For this reason, European policymakers, in close collaboration with technology experts and IT companies, started developing DCT standards based on the exchange of anonymized Bluetooth data. The European approach to DCT is defined in specific guidelines issued by the European Commission (EC) on April 17, 2020. This guidance is centered around the principle of data minimization, including precisely defined limits for data disclosure, use, and storage(4).

Meanwhile, in mid-April, the eHealth Network (comprising representatives of authorities responsible for digital health in the 27 EU Member States plus Norway) published a common toolbox, specifying essential requirements for European DCT apps. This toolbox emphasized a preference for decentralized protocols which store anonymized proximity data exclusively on users' mobile phones, over protocols storing data on centralized servers that are run by national health authorities. In particular, echoing the opinion of the European General Data Protection Board, this guidance underscored decentralized approaches as better suited to "keep personal data processing to the absolute minimum," enhance citizens' willingness to download and use DCT apps, and prevent "risks of data breaches and cyberattacks" (5).

At this time, many European technology experts were still collaborating on a centralized protocol called the Pan-European Privacy-Preserving Proximity Tracing protocol (PEPP-PT). Ultimately, though, some members of the PEPP-PT project resigned from this consortium in order to form a new protocol (6). The privacy-preserving decentralized protocol (Decentralized Privacy-Preserving Proximity Tracing, or DP-3T for short) was developed by a number of European

academic institutions, in conjunction with the Swiss Federal Institutes of Technology (ETH Zurich and the EPFL of Lausanne).

In the meantime, Google and Apple released an application programming interface (API) to implement this protocol on Apple and Android mobile operating systems (Google/Apple Exposure Notification system, or GAEN for short). Most decentralized DCT systems in Europe, including the Swiss model, run on this protocol. Countries such as Germany and the UK used the centralized model initially, but adopted the decentralized scheme powered by Google and Apple for the final version of their national DCT apps, introduced on June 16 and September 24, 2020, respectively (7, 8).

At the time of writing, 19 of the 27 EU Member States plus Switzerland have created a national DCT app (9). Of these, only France and Hungary have opted for a centralized solution (10).

In this comparative study of national proximity tracing apps, we seek to characterize the European approach to DCT, and to examine its evolution between the first and second waves of SARS-CoV-2. Our analysis shows that European DCT systems, to some extent, are evolving to incorporate new features extending their capabilities beyond mere proximity tracing – a development that calls for careful scrutiny and adequate oversight.

9.3 Methods

In order to examine the evolution of the European DCT landscape, we collected information from primary sources about national DCT apps in the following countries: France, Germany, Ireland, Italy, the Netherlands, Switzerland, and the UK (including England, Wales, and Scotland).

We included DCT systems released between March and October 2020. All the systems we included in our analysis revolve around a smartphone app as their key implementation technology. For inclusion in our study, the language of the app had to be English, French, Italian, or German (languages spoken by the authors). For each app, we collected *Privacy Policy* and the *Terms of Use* documentation from the app itself or its associated website. When available, we also analyzed the “FAQ section” and “press release” documentation, which usually contain a series of questions and answers, as well as concise information about the app’s functionality and

data processing. A list of the primary sources analyzed is available as supplementary material. Each source was archived (on archive.org) as it appeared at the time of review.

From this documentation, we first extracted and recorded general information and technical features for each DCT app (via MS Excel). Next, we imported the retrieved documents into Nvivo for qualitative content analysis. Two researchers (AB and AF) inductively created analytic codes from the text until thematic saturation was achieved (11). Semantically similar codes were further grouped into themes and subthemes. Two researchers (AB and AF) coded the text independently and resolved any coding discrepancies through discussion.

For comparative purposes, we collected information from secondary sources about national DCT systems in Asia. A list of these sources is available as supplementary material.

9.4 Results

Common characteristics

Appendix 1 Chapter 9 provides a summary of select descriptive features for each included DCT app. A certain degree of similarity is evident across the analyzed DCT systems. For example, all of them were developed in public-private partnerships between the state (or national health authority), software development companies, and, at times, research institutions. Furthermore, all of the apps function on a voluntary basis, in order to safeguard individual freedom. Moreover, a strong focus on privacy preservation and data protection is a common feature of the European approach to DCT. However, not all countries use the same architecture to achieve this aim.

The majority of DCT apps rely on decentralized protocols. These apps operate with the privacy-preserving technology framework released by Google and Apple, which allows matching codes to be kept on the user's phone, and in the case of a positive test, fetches only an anonymized ID from a centralized database, in order to check for high risk contacts. Among the apps we analyzed, only the French *TousAntiCovid* adopts a centralized approach to data storage. To justify this decision, the French government argued that the Google/Apple system contradicts the digital sovereignty of the state and does not provide sufficient privacy safeguards, as sensitive data about positive cases, albeit encrypted, are accessed by users' apps (12). Moreover, as the

FAQ section of the French app specifies, "*the Government considers that protecting the health of the French people is a mission that is the exclusive responsibility of the State and not of private international actors*."[1] (13)

From a technical perspective, European DCT apps employ similar exposure parameters (two meters apart for fifteen minutes) to notify app users of a potentially dangerous contact. Taking a precautionary approach, the German *Corona-Warn-App* uses the most stringent exposure parameters, alerting a user who is within eight meters and for at least ten minutes from an individual with a confirmed SARS-CoV-2 infection. The French *TousAntiCovid* employs the least stringent criteria of one meter apart for fifteen minutes.

Despite differences in data storage locations across countries, we noted that data retention periods are consistent, both for randomly generated ID codes as well as temporary exposure codes. Randomly generated ID codes are generally stored for fourteen days, while positive exposure codes are kept for 14 (Ireland, Italy, France, Netherlands, Scotland, and the UK) or 21 (Germany and Switzerland) days.

All of the reviewed systems collect statistical data concerning the number of users who downloaded the app, the number of apps actually in use, the positive cases uploaded to the system, the number of alerts sent to users, and the functioning of the app (e.g. Bluetooth signal strength, success of the data exchange, and the time at which the data must be destroyed). Some apps such as *SwissCovid* (Switzerland), *Immuni* (Italy), and *Corona-Warn-App* (Germany) have dedicated web pages offering aggregate information on how the respective DCT systems are being used (14-16). The apps also collect metrics data for public health surveillance, such as the day, time, and duration of a contact; whether the infected user is asymptomatic; the first day of illness; and the date of testing. Countries may retain such anonymous data for epidemiological surveillance or research purposes, however retention periods vary across countries. In Italy, metric (i.e. aggregated statistical) data is kept until the end of the emergency, but no later than 31 December 2021 (a limit previously set to the end of 2020.) In Ireland, England, and Scotland, metric data are retained respectively for at least seven years, twenty years, and indefinitely.

[1] Translated by the authors.

The seamless functioning of national DCT apps across borders motivated the European Commission to create an EU-wide system called *getaway*, to enable interoperability and help break the chain of COVID-19 infection across borders. The *getaway* would allow users who have installed one DCT app to travel to another participating European country and still receive contact tracing alerts (16). So far, however countries with interoperable apps include only Croatia, Denmark, Italy, Ireland, Germany, Latvia, the Netherlands, Poland, and Spain.

Country-specific features

European DCT apps differ in three respects: what happens upon notification of a contact with a positive case; how positive test results are handled; and additional features beyond proximity tracing.

In all cases analyzed, DCT apps advise users on what to do upon notification of close contact with someone who has tested positive for SARS-CoV-2. Most apps give users instructions for how to self-isolate, register for testing, and contact health authorities if symptoms emerge. The Irish *COVID Tracker* app allows users to voluntarily add a phone number, which is shared with health authorities. In case of close contact, the user not only is alerted by the app, but also phoned by the health authority that provides information about next steps and eventually arrange a COVID-19 test.

Each country follows its own procedure for uploading a positive test result into the DCT system. In Scotland for instance, health authorities send an exposure code via SMS to users who tests positive. Users enter the code, active for 72 hours, into the *ProtectScotland* app. In France, the code is sent to users in a link via email, and via post as a QR code. Users must thus enter personal information (mobile number, email, address) in order to communicate the outcome of a positive test and trigger notification to other users. Other countries have chosen methods which avoid this requirement. Users of the German *Corona-Warn-App* can scan a QR code linked to test results, automatically activating the exposure code. In Switzerland, Italy, and the Netherlands, users must phone the health authority upon notification of a positive test result, in order to activate the exposure code.

Our qualitative assessment explored the evolution of DCT apps as one component of broader policy efforts intended to curb the economic and public health effects of the pandemic. The most

recently released DCT apps were introduced after the summer, when a second wave of SARS-CoV-2 was already apparent in most European countries. At this time, some national apps released features that went beyond simple proximity tracing (see Figure 9.1). For instance, the French *TousAntiCovid* (the successor to a previous app called *StopCovid*, which was downloaded by a mere 2.6 million people and therefore replace by *TousAntiCovi*d (17)) expanded its functionality, allowing users to enter their postal code to receive more granular information about the local epidemiological situation. Moreover, users of the French app can access a government website (*Depistage COVID-19*) with a map of open testing centers and their current waiting times (18). The *NHS Covid-19* app, available in England and Wales, offers COVID-19 risk estimates as well. When users enter their postcode, they receive a notification of risk-level (low, medium, high) based on aggregate COVID-19 case information available to local authorities in a given area (19).

A daily symptom checker is integrated into the Irish *COVID Tracker* app, alongside the contact tracing function. This feature enables users to receive personalized recommendations (e.g. self-quarantine, call their physician, request a COVID-19 test) in relation to any symptoms and their severity, and to demographic data voluntarily entered into the system. The French app allows users to connect to a similar symptom checker, hosted on a separate government website (20). The *NHS Covid-19* app, one of the latest to be released in October 2020, integrates a symptom checker tool alongside the option to order a COVID-19 test, via a link to the NHS Test and Trace website. Users can then receive results directly through the app. These new features qualify the app as a medical device, as they enable collection of health data and provide personalized health recommendations to users. The *SwissCovid* and *NHS Covid-19* apps are the only European DCT apps registered as Class I medical devices (21).

The *NHS Covid-19* app also functions as a countdown tool for self-isolation, and a check-in instrument when visiting public venues. The first function, which calculates the length of time a user should self-isolate, activates automatically when a user is notified of contact with a positive case. Based on the encounter date, the app recommends the user self-isolate for ten days, beginning with the last encounter with the infected person. The countdown tool is also activated when a user enters COVID-19 symptoms or a positive COVID-19 test result into the app. The countdown is initiated respectively on the day on which symptoms first appear or on the day of the test.

The check-in function allows users to scan a QR code when entering public spaces such as restaurants, bars, shops, cinemas, or religious centers (22, 23). Location data is stored on the user's phone for 21 days. Authorities cannot access this information unless users decide to make it available. NHS documentation explains that the check-in function enables users to record locations visited. App users can then decide to voluntarily disclose this information to contact tracers in case they receive a positive test result. Contact tracers routinely collect information from individuals who test positive (whether they use the NHS app or not) about places visited in the days prior to the test. While individuals have the right not to disclose recent locations that they visited, this information allows contact tracers to alert others who visited the same location. UK health authorities use this information also to assess the level of risk based on the aggregate number of coronavirus cases reported at a particular venue in a certain time period, together with the type of venue (e.g. its architecture). This activity enables health authorities to update the list of places considered to be risky. When public health officials identify a venue as "at risk," they add to a national reference list that is synchronized with the *NHS Covid-19* app. The app can thus issue an alert to users who have checked in at a risky venue. The tone of the alert message is calibrated according to the level of risk identified by the local health protection team. If risk is high, the user may be urged to call the health authority immediately. The alert does include information about the venue itself.

Governance and oversight

Privacy Policies and Terms of Use documents provide information concerning the ethical conditions and legal bases for treatment of personal data by the respective national DCT systems. This information is meant to lend legitimacy to DCT activities, and to reassure the public about legal compliance.

All privacy policy documents of EU member states also make reference to the General Data Protection Regulation (GDPR), particularly concerning the protection of the rights of data subjects ensured by this Europe-wide legislation. For example, the Irish *COVID Tracker* privacy policy declares that "*The app is voluntary to use and the legal basis for the processing of the data is consent – namely Article 6(1)(a) of the GDPR for the processing of personal data and Article 9(2)(a) of the GDPR for the processing of special categories of personal data, in this case health related data*." (24)

In some cases, data governance principles are also reported. The privacy notice of the Italian DCT app *Immuni* declares compliance with Articles 13-14 of EU GDPR and respect for the principles of privacy ("*Under no circumstances will the users' movements be tracked, thus excluding any form of geolocation.*"), purpose limitation, and data minimization ("*Only the data necessary to alert the users that they have been exposed to a risk of infection, as well as to enable the adoption of any prevention and healthcare measures, are collected*"). (25)

In Switzerland as well, DCT documentation provides the legal basis for the processing of collected data, referencing both existing and new, ad hoc provisions: *"The federal legislation on data protection is applicable to the data processing. In addition, the Data Protection Statement is in line with the Epidemics Act of 28 September 2012 (EpG; SR 818.101) and the Ordinance of 24 June 2020 on the Proximity Tracing System for the Coronavirus SARS-CoV-2 (VPTS; SR 818.101.25)."* (26)

These documents frequently mention the role national data protection authorities and various expert bodies played in the early assessment of DCT apps. The FAQ section of *TousAntiCovid*, for example, explains that before the launch of the app, a number of national advisory bodies was consulted on the question of digital tools and privacy protection. The *Conseil Scientifique COVID-19* came out in favor of the app, affirming the usefulness of digital tools in light of the updated "*Test, Alert, Protect*" strategy. Furthermore, the documentation notes the approval of the CNIL (*Commission Nationale Informatique et Libertés*, the French data protection authority), which was responsible for assessing whether adequate data protection measures were in place, both before and after the launch of the app (13).

Documentation from various countries describes the effort to engage a broader array of societal actors in the development of the DCT system. For example, the documentation of the *ProtectScotland* app states that "*The Scottish Government and the NHS Scotland have rigorous information governance process in place. From the early stages of the design of the app, a thorough consultation with relevant Scottish groups of interests and advocacy has taken place, including: The Health and Social Care (Scotland) Public Benet and Privacy Panel; The Scottish Privacy Forum; The Open Rights Group; The COVID-19 Data and Intelligence Network – Data ethics and public engagement subgroup; and representatives of the general public.*" (27) However, no details are provided as to public engagement initiatives for the rest of the DCT systems in our sample.

In all cases, the analyzed documentation offers information concerning accountability for the lawful and responsible handling of personal data. For example the German *Corona-Warn-App*'s privacy policy reports that the app *"is provided by the Robert Koch Institute […]. The RKI is also what is called the controller under data protection law, meaning it is responsible for the processing of App users' data. You [the user] can contact the RKI's data protection officer at the above address."*(28) Likewise, the Dutch *CoronaMelder* privacy policy cites the Minister of Health, Welfare and Sport, and the Regional Health Service, as controllers and accountable bodies for the protection of user data against potential abuse, loss, unauthorized access, unwanted disclosures, and unauthorized changes (29). Our study indicates that national governments and departments of health are the authorities responsible for the good functioning of DCT apps, as well as for communication with users and/or intervention when issues arise.

Despite efforts towards the transparent governance of DCT apps, limited information is available about oversight bodies and mechanisms charged with regularly assessing the functioning of DCT systems.

Two exceptions are the commitment by the NHS in England and Wales to review the privacy impact assessment in the event of software updates. As mentioned previously, this app "*is CE marked as Class I medical device in the United Kingdom and developed in compliance with European Commission Directive 93/42/EEC for Class I devices*." (23) As such, the app is subject to stricter oversight regulation (30). The Scottish DCT app also provides some details about the oversight mechanism in place; its documentation states that "*any future changes [to the app] will follow rigorous scrutiny; the decision will be balanced against public health benefit and cost (balanced against other health priorities) and this privacy notice will be updated accordingly for transparency*." (27)

Apart from these two cases, DCT documents do not relay how the responsible institutions intend to monitor an app's activity and the addition of new features over time. Notably, the Dutch documentation stresses that it is the responsibility of the user to check for data information notice updates (which may be introduced with future app developments). These changes will be in immediate effect in the app following publication of the updated privacy policy. Similarly, all of the *Terms of Use* that we analyzed encourage users themselves to inspect the app's source code

(via online platforms such as GitHub/GitLab), as well as to report back about their experience of using the app (including any potential issues).

9.5 Discussion

The European approach to DCT has been characterized by marked attention to privacy preservation and data protection. The General Data Protection Regulation (GDPR), in force since May 2018 in European Member States, played a central role in shaping this approach. The GDPR considers the protection of natural persons in relation to personal data processing as a fundamental right (rec. 1 GDPR). Moreover, it recognizes the challenges that new technological developments, together with the global reach of big technology corporations, pose to the protection of personal data (rec. 6 GDPR). Article 25 of the GPDR espouses the principles of data protection by design and by default, making them a legal requirement. These requirements arguably played a key role in shaping the European approach to DCT.

In general, data protection by design asserts that data processing activities should adopt state-of-the-art data protection safeguards across all technical components and processes. Data protection by default refers to the principle that data processing options should automatically be set to the most privacy preserving mode. From a practical point of view, these principles translate into a series of requirements, including data minimization and individual control of personal data. Data minimization contends that only data strictly necessary for a specific purpose should be collected and used, and there must be fixed limits on the extent of processing and the duration of storage and accessibility (art 25.2). Individual control refers to the principle that personal data should be made accessible only upon authorization of data subjects.

These provisions ensure the voluntary nature of European DCT systems, and the selection of privacy-preserving technological solutions for DCT. In particular, the use of GPS-based DCT was never given consideration in Europe, as all countries surveyed recognize Bluetooth-based models as the only legally viable option. In some countries, such as Italy for example, technology experts did not rule out *a priori* the possibility of collecting limited amounts of geolocation data for DCT purposes, but this option never gained support in policy circles. The rationale, based on data protection by design, is that geolocation data is considered redundant to the aim of proximity tracing, since it contains more information than is necessary to notify users about contact with

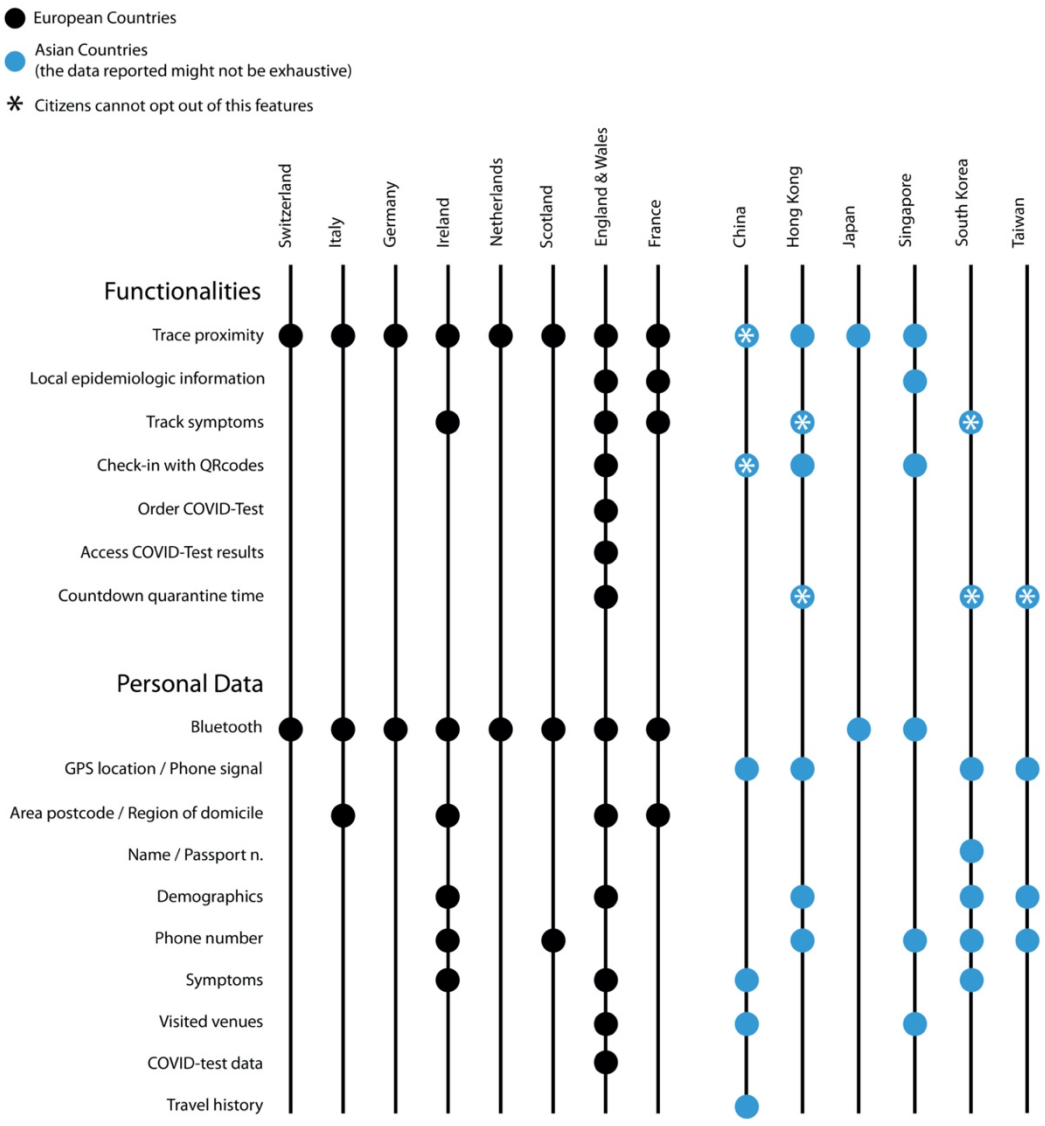
positive cases. However, this argument depends upon a specific view of DCT as a personal warning system, rather than a public health surveillance tool.

European policymakers and advisors however showed a lesser degree of consensus as to the best IT architecture for DCT systems. In the view of some stakeholders, the GDPR did not appear to pose a concrete constraint on specific technological options for DCT. Germany and the UK initially favored a centralized model, to later change to a decentralized one. France and Hungary (not reviewed) ultimately implemented centralized DCT, while remaining fully compliant with GDPR rules. Switzerland, while not a member of the European Union, is revising its Federal Act on Data Protection (FADP) in a way that will also ensure general alignment with the provisions of the GDPR, especially regarding the rights of data subjects. The newly approved law (expected to come into effect in 2022) endorses privacy by design and also by default.

The European model differs in meaningful ways from the DCT approaches adopted by Asian countries in the earliest phases of the pandemic. While it is not possible to speak of an 'Asian model' due to the great diversity among DCT systems in Asian countries, it is evident that a more expansive approach characterizes DCT in countries such as China, Hong Kong, Singapore, South Korea, and Taiwan (see Figure 9.1).

DCT apps developed in China at the beginning of the pandemic became mandatory immediately (31). Hong Kong, Taiwan, and South Korea also deployed mandatory apps and wearable trackers for those living under quarantine, either due to testing positive for COVID-19 or returning from foreign travel (32-34). These apps record GPS geolocation data or use cell tower data to ensure that individuals remain in their homes while in quarantine, and ask the user to enter symptoms, in order to monitor the course of the disease. Taiwan for example used the quarantine DCT feature in combination with rigorous manual contact tracing, which successfully helped contain the spread of the disease (35). South Korea relied on more intrusive surveillance measures, including a number of system tracking citizens' movement, and interactive maps displaying locations visited by COVID-19 positive individuals (36). Singapore was one of the first countries worldwide to introduce a voluntary centralized digital contact tracing app called *TraceTogether*, which was later integrated with a check-in system (called *Safe Entry*) for entry into public spaces (mandatory from the beginning of January 2021) (37). A similar feature was adopted in October 2020 in Hong Kong, where the government is still debating whether the app will be made mandatory.

Figure 9.1: Features of DCT systems in selected European and Asian countries: functionalities and types of personal data collected (Updated to 31.10.2020)



During the period examined in this study (March to October 2020) European DCT systems showed stability in their overall technical architecture. In all of the reviewed countries, participation

in DCT was originally designed to be, and remained, entirely voluntary. Data collection remained limited to randomly generated and periodically deleted Bluetooth IDs. While organizational and technical improvements were implemented to streamline the uploading of positive test results, this process remained fully voluntary, with disclosure of test results possible only with explicit authorization by a DCT user. One partial exception is presented by the England & Wales app, which automatically uploads test results when a test is booked directly through the app.
We have observed an expansion of DCT app features beyond basic proximity tracing in European apps released or updated during the second wave of SARS-CoV-2 that swept through Europe during late summer 2020. Novel features include the capability to track symptoms (Ireland, France, England, and Wales), acquire more detailed epidemiological information about a given area (France, England, and Wales), check in at venues (England and Wales), order COVID-19 tests and access results (England and Wales), and count down the quarantine time (England and Wales).

Our study indicates that privacy preservation through state-of-the-art technological solutions and alignment with data protection laws is the key defining feature of the European approach to DCT. However, we also demonstrated how such an approach is evolving to incorporate novel technological capabilities beyond mere proximity tracing. These developments signal a shift from a strict interpretation of data minimization and purpose limitation, towards a more expansive approach to digital contact tracing in Europe. This evolutionary trajectory seems to reflect technological capacities already seen in Asian countries.

In Europe, the incorporation of novel capacities seems a response to two aims. On one hand, adding features can be viewed as a way to encourage users to download and use DCT apps by offering additional functionalities that users may find useful or interesting. Considering the relatively low level of uptake of DCT apps in European countries compared with the adoption rates needed to ensure effectiveness (38), adding new features may be seen as one way to deliver more personal utility to app users, thus incentivizing participation. On the other hand, novel features such as digital check-ins may increase the aggregate data available to public health authorities, expanding their capacity to monitor how the epidemic is evolving and how the population responds to containment measures. Furthermore, this feature is an ingenious way to integrate manual and digital contact tracing. Both manually and digitally collected information about the whereabouts of positive cases can contribute to map out risky locations. In turn, this information can be used to alert people about potential contacts with positive cases irrespective

of whether the use the DCT app or not, thus extending the utility of the DCT app beyond the section of the population that is actually using it.

The panorama of European DCT systems is evolving also in other respects. In December 2020, the privacy policy of the *Corona-Warn-App* was updated, allowing users to record symptoms and retrieve test results (39). In France, the government is considering adding a check-in function to the *TousAntiCovid* app when reopening restaurants (40). These updates may prelude to further expansion of DCT app capabilities in the near future. In Italy for example, the possibility of using the *Immuni* app as a tool in the imminent vaccination campaign is being discussed. The app could evolve into a digital booking system for vaccination appointments, and could then be licensed to store a digital copy of the vaccination certificate for display to health authorities, for entry to designated places or activities (41).

The possible evolution of European DCT systems calls for careful scrutiny and appropriate oversight, especially with respect to GDPR provisions. It must be noted that novel features do not necessarily contravene the principle of data minimization, as they can still be based on minimum necessary data collection for data processing purposes. However, such new features expand the scope of DCT apps beyond the purpose of proximity tracing and warnings to individual users. The legally mandated safeguards regarding data collection and storage may therefore be insufficient to capture additional privacy risks linked to novel functionalities. In other words, data protection by design and by default may be inadequate to address the evolution of DCT systems. To be sure, DCT innovation does not necessarily create greater privacy risks. Such technological evolution should not be prevented, and both public health and ethical rationale support changes aimed at improving the effectiveness of DCT systems against the spread of the virus. Yet as the purpose of DCT apps expands to incorporate new capacities, privacy risks should be regularly reassessed. An adaptive governance approach to DCT seems best suited to regularly fine tuning governance structures and oversight mechanisms over time (42), thus capturing the technical evolution of such systems and their ethical, legal and societal implications.

9.6 Conclusion

As they face subsequent epidemic waves, European countries are tasked with deploying all possible means to mitigate the spread of the virus and minimize the health-related, personal,

economic, and social damage it has caused since early 2020. Digital methods offer valuable aid to contain this disaster. In the context of harsh measures and restrictions to individual freedom made necessary by the emergency, DCT is relatively more tolerable, especially in its European incarnation, which offers a comprehensive set of technical and legal safeguards against potential abuse of personal data. Nevertheless, the vast majority of European citizens have not downloaded national DCT apps, despite their ethical and technical robustness. It remains unclear whether this is due to insufficient trust or to a perceived lack of personal benefit associated with use of these apps.

In this study we reviewed DCT systems in a number of European countries. We highlighted the strong emphasis that all such systems place on privacy and data protection, their fully voluntary character, and their adoption of the same Bluetooth-based standards for proximity tracing. We noted that such ethical and technological commitment is enshrined in both centralized and decentralized DCT systems. Furthermore, we reported an emerging evolutionary trajectory resulting in the incorporation of novel technological features beyond mere contact tracing that are, to some extent, reminiscent of those already seen in Asia. However, additional policy efforts seem necessary to account for such developments, to gain public trust and to foster more widespread adoption of DCT as a valuable means for containing the effects of the SARS-CoV-2 pandemic.

Ethics statement

Not applicable

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author Contributions

AB and AF contributed to data collection, analysis, manuscript drafting and editing. EF contributed to writing and editing the manuscript. All authors made a substantial intellectual contribution to the manuscript and approved the submitted version.

Funding

This manuscript was supported by the Swiss National Science Foundation under awards NRP75 (# 407540_167223) and NRP77 (#407740_187356).

References

1. Johns Hopkins. COVID-19 Map - Johns Hopkins Coronavirus Resource Center (2021) [Accessed 27.01.2021]. Available from: https://coronavirus.jhu.edu/map.html.
2. European Centre for Disease Prevention and Control. COVID-19 situation update for the EU/EEA, as of week 2 2021 (2021) [Accessed 27.01.2021]. Available from: https://www.ecdc.europa.eu/en/cases-2019-ncov-eueea.
3. Han E, Tan MMJ, Turk E, Sridhar D, Leung GM, Shibuya K, et al. Lessons learnt from easing COVID-19 restrictions: an analysis of countries and regions in Asia Pacific and Europe. *The Lancet* (2020). doi: https://doi.org/10.1016/S0140-6736(20)32007-9.
4. EUR-Lex. Guidance on Apps supporting the fight against COVID 19 pandemic in relation to data protection (2020/C 124 I/01) (2020) [Accessed 27.01.2021]. Available from: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020XC0417%2808%29.
5. eHealth Network. Interoperability guidelines for approved contact tracing mobile applications in the EU (2020) [Accessed 27.01.2021]. Available from: https://ec.europa.eu/health/sites/health/files/ehealth/docs/contacttracing_mobileapps_guidelines_en.pdf.
6. NS TECH. PEPP-PT vs DP-3T: The coronavirus contact tracing privacy debate kicks up another gear (2020) [Accessed 27.01.2021]. Available from: https://tech.newstatesman.com/security/pepp-pt-vs-dp-3t-the-coronavirus-contact-tracing-privacy-debate-kicks-up-another-gear.
7. Reuters. Germany flips to Apple-Google approach on smartphone contact tracing (2020) [Accessed 27.01.2021]. Available from: https://www.reuters.com/article/us-health-coronavirus-europe-tech-idUSKCN22807J.
8. Medical Device Network. The UK's contact-tracing apps: why the long wait? (2020) [Accessed 27.01.2021]. Available from: https://www.medicaldevice-network.com/features/uk-contact-tracing-app/.
9. European Commission. Mobile contact tracing apps in EU Member States (2020) [Accessed 27.01.21]. Available from: https://ec.europa.eu/info/live-work-travel-eu/coronavirus-response/travel-during-coronavirus-pandemic/mobile-contact-tracing-apps-eu-member-states_en.
10. Henrich Böll Stiftung. Pan-European contact-tracing – a chance for more normality and freer movement | Heinrich Böll Stiftung | Brussels office - European Union (2020) [Accessed 27.01.2021]. Available from: https://eu.boell.org/en/2020/11/16/pan-european-contact-tracing-chance-more-normality-and-freer-movement.
11. Ando H, Cousins R, Young C. Achieving saturation in thematic analysis: Development and refinement of a codebook. *Comprehensive Psychology* (2014) 3:03. CP. 3.4. doi: https://doi.org/10.2466/03.CP.3.4.
12. Politico. How Google and Apple outflanked governments in the race to build coronavirus apps (2020) [updated 2020-05-15Accessed 27.01.2021]. Available from: https://www.politico.eu/article/google-apple-coronavirus-app-privacy-uk-france-germany/.

13. Ministère des solidarités et de la santé. TousAntiCovid: Réponses à vos questions (2020) [Accessed 31.10.2020]. Available from: https://web.archive.org/web/20201031233234/https://solidarites-sante.gouv.fr/soins-et-maladies/maladies/maladies-infectieuses/coronavirus/tousanticovid.
14. Federal Statistical Office. SwissCovid App Monitoring (2021) [Accessed 27.01.2021]. Available from: https://www.experimental.bfs.admin.ch/expstat/en/home/innovative-methoden/swisscovid-app-monitoring.html.
15. Immuni Italia. The numbers of Immuni (2021) [Accessed 27.01.2021]. Available from: http://www.immuni.italia.it.
16. Robert Koch Institut. Kennzahalen zur Corona-Warn-App (2021) [Accessed 27.01.2021]. Available from: https://www.coronawarn.app/assets/documents/2021-01-22-cwa-daten-fakten.pdf.
17. The Connexion France. France to release 'more interactive' StopCovid app (2020) [Accessed 27.01.2021]. Available from: https://www.connexionfrance.com/French-news/France-is-to-release-more-interactive-StopCovid-app-on-October-22.
18. Santé.fr. DepistageCovid (2020) [Accessed 31.10.2020]. Available from: http://web.archive.org/web/20201031215938/https://sante.fr/recherche/trouver/DepistageCovid.
19. GOV.UK. DPIA annex 2: high-risk postcode district user journey flow (2020) [Accessed 27.01.2021]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/920835/HighRiskPostcode_20200923.pdf.
20. Ministère des solidarités et de la santé. Mes Conseils Covid. Des conseils personnalisés pour agir contre le virus (2020) [Accessed 27.01.2021]. Available from: https://mesconseilscovid.sante.gouv.fr/.
21. Decomplix. Decomplix supports FOPH with SwissCovid App for contact tracing (2020) [updated 2020-06-25Accessed 27.01.2021]. Available from: https://decomplix.com/swisscovid-app-contact-tracing/.
22. NHS COVID-19 app support. FAQ NHS Covid-app: Which venues in England should display the official NHS QR code poster? (2020) [Accessed 27.01.2021]. Available from: https://faq.covid19.nhs.uk/article/KA-01183/en-us?parentid=CAT-01043&rootid=CAT-01027.
23. NHS COVID-19 app support. NHS COVID-19 app: Terms of use (2020) [Accessed 31.10.2020]. Available from: https://web.archive.org/web/20201030072744/https://covid19.nhs.uk/our-policies.html.
24. Health Service Executive. COVID Tracker app: Data Protection Information Notice (2020) [Accessed 31.10.2020]. Available from: https://web.archive.org/web/20201004220235mp_/https://covidtracker.gov.ie/privacy-and-data/data-protection/#6.
25. Immuni Italia. Immuni: Privacy Notice (2020) [Accessed 31.10.2020]. Available from: https://web.archive.org/web/20200805213934/https://www.immuni.italia.it/app-pn.html.
26. Swiss Federal Office of Public Health. SwissCovid App: Data Protection Statement (2020) [Accessed 31.10.2020]. Available from: https://web.archive.org/web/20201026031819/https://www.bag.admin.ch/dam/bag/en/dokumente/cc/kom/swisscovid-app-datenschutz.pdf.download.pdf/FOPH_SwissCovid_Data_Protection_Statement_24_June2020.pdf.
27. NHS Scotland. ProtectScotland: How we use your data (2020) [Accessed 31.10.2020]. Available from: https://web.archive.org/web/20201030192549if_/https://www.protect.scot/how-we-use-your-data.

28. Robert Koch Institute. Corona-Warn-App: Privacy notice (2020) [Accessed 31.10.2020]. Available from: https://web.archive.org/web/20201130083701/https://www.coronawarn.app/assets/documents/cwa-privacy-notice-1.3-en.pdf.
29. CoronaMelder. CoronaMelder: Privacy Statement (2020) [Accessed 31.10.2020]. Available from: https://web.archive.org/web/20201015200129/https://coronamelder.nl/en/privacy.
30. Pane J, Francisca RD, Verhamme KM, Orozco M, Viroux H, Rebollo I, et al. EU postmarket surveillance plans for medical devices. *Pharmacoepidemiology and drug safety* (2019) 28(9):1155-65. doi: http://dx.doi.org/10.1002/pds.4859.
31. CNN Business. China is fighting the coronavirus with a digital QR code. Here's how it works (2020) [Accessed 27.01.2021]. Available from: https://www.cnn.com/2020/04/15/asia/china-coronavirus-qr-code-intl-hnk/index.html.
32. Quartz. Hong Kong is using tracker wristbands to geofence people under coronavirus quarantine (2020) [Accessed 27.01.2021]. Available from: https://qz.com/1822215/hong-kong-uses-tracking-wristbands-for-coronavirus-quarantine/.
33. MIT Technology Review. South Korea is watching quarantined citizens with a smartphone app (2020) [Accessed 27.01.2021]. Available from: https://www.technologyreview.com/2020/03/06/905459/coronavirus-south-korea-smartphone-app-quarantine/.
34. Quartz. How Taiwan is tracking 55,000 people under home quarantine in real time (2020) [Accessed 27.01.2021]. Available from: https://qz.com/1825997/taiwan-phone-tracking-system-monitors-55000-under-coronavirus-quarantine/.
35. Cheng H-Y, Jian S-W, Liu D-P, Ng T-C, Huang W-T, Lin H-H. Contact tracing assessment of COVID-19 transmission dynamics in Taiwan and risk at different exposure periods before and after symptom onset. *JAMA internal medicine* (2020) 180(9):1156-63. doi: 10.1001/jamainternmed.2020.2072.
36. CNN Business. Coronavirus mobile apps are surging in popularity in South Korea (2020) [Accessed 27.01.2021]. Available from: https://www.cnn.com/2020/02/28/tech/korea-coronavirus-tracking-apps/index.html.
37. SafeEntry. Places where SafeEntry must be deployed (2021) [Accessed 27.01.2021]. Available from: https://safeentry.gov.sg/.
38. Toussaert S. Upping uptake of COVID contact tracing apps. *Nature Human Behaviour* (2021). doi: 10.1038/s41562-021-01048-1.
39. German Federal Government. Corona- Warn-App Privacy notice (2020) [Accessed 27.01.2021]. Available from: https://www.coronawarn.app/assets/documents/cwa-privacy-notice-en.pdf.
40. La Depeche. TousAntiCovid : de nouvelles fonctionnalités annoncées pour la future "réouverture des restaurants" (2020) [Accessed 27.01.2021]. Available from: https://www.ladepeche.fr/2020/11/23/tousanticovid-de-nouvelles-fonctionnalites-annoncees-pour-la-future-reouverture-des-restaurants-9216559.php.
41. La Repubblica. Immuni, la seconda vita: "Così la app può dialogare con il piano vaccini" (2020) [Accessed 27.01.2021]. Available from: https://rep.repubblica.it/pwa/generale/2020/12/18/news/immuni_la_seconda_vita_cosi_la_app_puo_dialogare_con_il_piano_vaccini_-278976650/.
42. Blasimme A, Vayena E. What's next for COVID-19 apps? Governance and oversight. *Science* (2020) 370(6518):760-2. doi: 10.1126/science.abd9006.

SECTION III

POLICY DIRECTIONS
AND THE FUTURE OF BIG DATA ETHICS

CHAPTER 10: General Discussion

Through various conceptual and empirical studies, this thesis examines the ethical considerations of big data in health research and health apps. Furthermore, it explores the consequences of these ethical considerations for existing ethical oversight mechanisms and governance.

Although the potential of big data for health research and digital health applications is substantial, new challenges arise from the collection, use, and analysis of copious amounts of user-generated and passively collected data. These challenges have ethical implications which, if not addressed by the regulations governing big data usage, have the potential to cause harm and eliminate opportunities to improve public health. Therefore, this thesis considers whether the ethical guidance that regulates health research and health apps is up-to-date and adequate to tackle the ethical concerns of big data and AI-enabled technologies. The presence of regulatory gaps, as well as fragmented guidance increases uncertainty among stakeholders. Consequently, uncertain stakeholders may overlook important ethical issues and choose solutions that encourage data usage that is detrimental to public interests and collective values.

As such, this thesis attempts to identify the gaps in the existing ethical guidance by studying the ethical oversight mechanisms present in health research and health apps. Once individual gaps are identified, this research addresses them by promoting tailored policies and reforms. Particularly, this thesis offers guidance to improve research ethics review processes, as well as recommendations to streamline regulatory approaches to health apps. The successful definition of effective governance and ethical oversight may ensure ethically aligned decision-making across relevant stakeholders and promote health benefits.

10.1. Main findings

This section summarizes the main findings of the scientific contributions presented in Chapters 2-9 and discusses them in a broader scientific context. Please refer to the results and discussion sections of each single chapter for a more detailed analysis of issues.

SECTION I: Key ethical issues in big data uses and AI-enabled technologies

Our findings highlight the research community's growing attention towards the use of big data in the health sector. Many authors discuss the promise that big data holds for improving individual and collective health, as well as health services (1, 2). However, the literature reviewed in Chapter

2 also highlights the many technical and methodological challenges arising at various stages of the data lifecycle, when collecting, handling, combining, and analyzing data. Not surprisingly, researchers using health data focus primarily on challenges such as data quality, integrity, and representativeness. Indeed, these technical issues undermine not only the usability of the data, in the case of corrupted, incompatible, or non-interoperable records, but also the reliability of the data analysis. Poor outcomes of the data analysis, in turn, undermine the validity of studies and the effectiveness of digital tools (3).

As our thematic analysis of the literature illustrates, the data quality is fundamental to ensure the accuracy of the AI algorithms. AI systems need big datasets to train on, to recognize statistically relevant correlations within new datasets (4). In this respect, our results show two problems that are two sides of the same coin. On one side, there is data representativeness. If the data provided to an AI system are not sufficiently representative of the population in which the algorithm is used, then the outcomes will be biased and unreliable for some individuals within that population group (5). For example, a study showed that artificial intelligence makes many errors in identifying potential melanomas on dark skin when trained on clinical sample images of light skin (6). The use of data from predominantly white adults and males of European or North American origin represents a fundamental problem in biomedical and health research (7). Such biased datasets make it overly complex and potentially risky to generalize research findings and AI outcomes to minorities or other population groups (distinguished by age, socioeconomic status, or geographical origin). On the other side, there is data accuracy. Given the widespread use of digital technologies, users can independently report their data in real time. However, user-generated data can introduce biases in the datasets due to their heterogeneity and imprecision (8, 9). Our findings emphasize that using inaccurate data to train an AI system can result in outputs which reproduce and automate those same inaccurate features. AI's inability to self-assess the reliability of its own outcomes can exacerbate this risk, potentially becoming a real danger in the health sector, where people's lives are at stake.

Another debated technical issue is transparency in AI systems applied to the health sector (Chapter 3). The growing attention to this topic comes from the GDPR provisions of data protection and data process transparency (10). The fact is that ML opacity brings along many ethical questions about the reliability of algorithmic predictions and correlations (11). The scientific validation of ML outcomes is much more complex when the machine processes cannot be explained in causal terms (12), directly impacting the ethics of doctor-patient relationships (13).

Our conceptual analysis highlights that physicians using opaque AI models rely on a tool that they cannot understand and for which no causal explanation can be given. This undermines the decision-making role of doctors when they have no firm basis on which to make decisions. This lack of reasoning may consequently reduce patient trust in doctors and the healthcare system. In this respect, the literature questions whether and under what circumstances patients should be informed about the use of AI in a medical context (14). Finally, algorithm opacity also has problems of legal liability - especially in cases of algorithm prediction error and medical negligence (15).

These findings illustrate a connection between technical limitations, such as data bias and algorithmic opacity, and a variety of ethical consequences, such as violations of the principle of justice, potential group level discriminations, reduced trust in healthcare providers and health technology, and risk of harm to individuals (16). These ethical concerns, however, are not prominently debated in the field of health research and digital health applications. As our literature reviews illustrate, the ethical discussion is drastically tilted towards privacy issues with researchers and app developers mainly focusing on compliance with existing data regulations. As of 2018, the GDPR provides the European legal framework for the collection and use of personal data. It guarantees a set of rights for each data subject (such as the right to know what personal data a given organization has collected, for what reasons, how it was obtained, and with whom it is shared) which are matched by obligations for researchers and developers (17). However, in other countries around the world, different sets of rules apply (Chapter 4). To compensate for the absence of minimum global privacy standards, the big data research community has been relying on robust technical solutions to store data safely, to prevent malicious actors from violating the privacy of individuals, and to allow data to be shared and used without identifying the data subjects (18). “Privacy by design” is the approach whereby privacy protections are embedded within the digital health technologies themselves (19). Nevertheless, the practical application of this concept is difficult. There is still an unresolved question of what the data subjects’ expectations are concerning their privacy and confidentiality in various contexts (e.g., data used by digital health applications vs. hospitals, or academic vs. commercial use of data) (20). At present, biomedical and health research most often requires an informed consent (as highlighted in Chapter 2). However, if the data are collected through health apps and commercial platforms, data processing is only bound by a cursory approval of the terms of service (21, 22). Although the challenge of privacy is key, this thesis – in accordance with previous literature (23) – argues that it should not monopolize the ethical debate. Therefore, previously

underexplored considerations of data ownership and control, research accountability, civic empowerment, technology accessibility, fair distribution of risks and benefits, and group-level harm mitigation should play a role in setting the rules to ethically collect and use big data and AI-enabled technology.

SECTION II (A): The state of ethical oversight in health research: existing gaps and key reforms

This thesis shows that interest in, and willingness to regulate the ethical use of big data and AI-enabled technology in health research is growing due to new technical and methodological challenges. Our findings reveal that regulators and other stakeholders published an increasing number of diverse soft-law and policy documents in this field (Chapter 5). However, the inconsistency of recommendations offered in these documents reflects a fragmented ethical and regulatory landscape. Indeed, there is no internationally accepted framework for conducting ethically aligned research. In addition, only a minimal fraction of these documents directly addresses research oversight mechanisms and their duties. Instead, most of the regulatory recommendations concerning the ethical, legal, and social implications of big data for research do not clarify which actors must promote and enforce these recommendations.

Section II A indicated that the absence of a uniform approach to tackle the ethical implications of big data research leaves RECs uncertain. This uncertainty is also apparent from our qualitative study (Chapter 6). Although the Swiss RECs agreed that big data in medical research has its problems, they disagreed on many aspects such as identifying the most pressing ethical implications and how to address them. In addition, some of the interviewees were alarmed by the shortcomings of the current model of ethics review considering big data novelties, while others firmly rejected any big data and AI exceptionalism in health research when compared to the risks brought by traditional biomedical research. This variety of perspectives makes the practice of ethics review inconsistent within and across countries if left without uniform normative guidance (24, 25). This could result in the case where one REC deems a big data project ethically aligned when another would judge it as inappropriate.

Our analysis of the current state of RECs shows several weaknesses when considering big data research. These weaknesses make RECs inadequate to fulfil their role successfully. Alongside the traditional weaknesses of RECs, Chapter 7 maps RECs' weaknesses specific to big data research, splitting these novel weaknesses in two categories through a conceptual analysis. The

first category concerns the RECs' scope. The results, in accordance with previous research, show that the regulations governing research ethics often allow big data projects to evade RECs' assessment (26). In fact, studies involving the use of anonymously collected data or conducted outside academia by private actors and citizen science groups do not necessarily fall under the purview of RECs. As our findings show, the problem lies in the fact that a vacuum in the ethical oversight guidance scheme may incentivize self-regulatory approaches. Bypassing ethics review can pose risks of harm to individual data subjects and society, while potentially reducing public trust in science (27). The second category deals with RECs' functional weaknesses. For instance, RECs can fail to appropriately review big data projects because they lack adequate expertise in data analytics and computer science. Furthermore, RECs rely on inadequate ethical tools to assess big data projects such as the traditional informed consent form, a definition of risk focused solely on the individual, and the ex-ante review process. These tools are unfit for this purpose, as big data research differs in many ways from traditional biomedical research (Chapter 7). As an example, big data research tries to learn from the data without necessarily formulating clear research questions. Moreover, big data analytics introduces unforeseeable risks which cannot be fully assessed at the beginning of the research, when the research protocol is submitted for review to the REC.

Despite all the above-mentioned profound weaknesses, this thesis stresses that RECs can still play a critical role as a key oversight mechanism for research. However, to successfully tackle the incipient challenges of health research with big data, REC mechanisms need reform. Chapter 7 provides a detailed explanation of how to improve the ethics oversight process. On one hand, the proposed reforms align with the available recommendations for RECs (Chapter 5), and on the other hand, they aim to address RECs' needs and requests (Chapter 6). We therefore present reforms at two levels: at the level of regulations and of procedures. Concerning the former, this thesis suggests reforming the existing research ethics legislation to expand RECs' purview beyond the current mandate, as well as to formulate new guidance on big data research, and to harmonize these standards across countries. Considering the latter, RECs' members should improve their expertise (through trainings and capacity-building strategies), diversify their composition (by including computer scientists or data analytics experts), and engage more with the stakeholders. This thesis emphasizes the need for an oversight mechanism that is flexible and adaptive to the challenges posed by big data research and the rapidly changing world of digital health. We therefore reflect on the possibility of improving the current oversight mechanisms by introducing complementary forms of oversight (in the form of "data review

boards"). However, simply increasing the number of committees could jeopardize the ethics review process by augmenting redundancy and inconsistency. Therefore, if complementary forms of oversight are introduced, they must be coordinated within a shared regulatory framework that defines best practices and clear assessment criteria.

SECTION II (B): The state of ethical oversight in health apps: existing gaps and key reforms

Similar to the context of health research, the use of health apps also prompts reflection about existing ethical guidelines and the oversight tools in place to regulate them. This thesis sheds light on two main gaps in the guidance of health apps.

First, health app guidance is fragmented and not comprehensive. As Chapter 8 shows, various stakeholders define their own regulatory standards, which focus on some ethical issues while disregarding others. For example, data protection authorities tend to focus on privacy and data security, health authorities on clinical validity and safety, private companies on accountability and legitimacy, and consumer groups on benefit distribution and public engagement. Moreover, the standards for assessing app quality not only variate across jurisdictions (28) but also within countries (29). This jungle of scattered regulations leaves stakeholders struggling to orient themselves and emerging harms unaddressed. This problem is also exacerbated by the lack of clarity of ethical standards (30), as app developers often do not know how to translate high level ethical principles into operationalizable practices.

Second, unclear governance is compounded by the weak role of oversight mechanisms. The literature clearly shows that most commercially available health apps not classified as medical devices, are not vetted by oversight mechanisms, despite potentially being of low quality and risky to users (31). This normative difference is also seen in Chapter 9, as the digital contact tracing apps for COVID-19 released in England and Switzerland – because they are registered as medical devices – were subject to more stringent ethical scrutiny than in other countries. Oversight mechanisms for digital health applications also suffer from problems of technical expertise (32). As depicted by our results, health authorities often oversee health apps, even when they are unable to assess whether standards for data security, confidentiality, accuracy, and validity are sufficiently robust. An additional obstacle for ethical oversight uncovered in this thesis is how to stay effective despite the ever-changing nature of digital health apps. New app features and updates do not necessarily contravene data privacy and security provisions,

however, as apps become more sophisticated, there may be a greater variety of data collected and purposes for which they are used. As a result, additional risks related to new features - beyond just those of privacy - may arise (Chapter 9).

Finally, the rapidly changing landscape of health apps and the lack of clear ethical guidance presented in Section II B calls for interventions with targeted policies. Some authors propose an adaptive governance for digital health apps which would improve the effectiveness of oversight mechanism over time by addressing technological evolutions and their unforeseen ethical, legal, and social implications (33). In addition, the results of this thesis suggest the need for a streamlined governance process for health apps. That is, legal requirements for data uses should be clearer and more homogeneous given the purposes for which the data are used, regardless of their origin, the stakeholders who use them, or the jurisdiction in which they are used. This new governance should aim to promote the common good, thus integrating shared values such as just distribution of benefits and fair access to health services, while safeguarding individual rights.

10.2. Limitations and future research

This thesis has several limitations, both content-related and methodological. Regarding the former, the rapidly evolving field of big data and AI in healthcare is a primary limiting factor. This research suggests that the variety of big data uses and applications is growing rapidly over time. Therefore, any attempt to provide an up-to-date overview and comprehensive assessment - going beyond a specific historical moment - is inevitably limited. To reduce this shortcoming, it will be necessary to constantly update research on this topic. For example, future research should consider periodically replicating the various overview studies that are part of this thesis, integrating any emerging technological novelties. The methodology used in Chapters 2,5, and 8 offers a solid basis for replicating such studies (see Appendices Chapter 2,5,8).

Furthermore, the content of this thesis is thematically limited. It focuses exclusively on the ethical and governance implications of big data in the context of medical research and digital health applications (specifically health apps). These contexts were selected due to their relevance to various stakeholders, and because they constitute grey areas of the current governance landscape. However, as mentioned in Chapter 1, the field of digital health and medicine is broad, meaning that many of the issues discussed in this thesis could be valuable – subject to

appropriate adjustment and adaptation – to the deployment of other biomedical technologies and health tools (such as medical apps, wearable and implantable technologies, robotics, brain-computer interfaces, and HIT.). In fact, a growing body of literature explores these exact technologies and tools (34), thus compensating for this thesis' thematic limitation. Future research should also consider whether the ethical issues raised in this thesis can be generalized beyond the health sector to other fields, such as the labor market or banking services. In fact, the digital economy creates an interconnected system of data where decisions made in one area can have spill-over effects in another. This creates a need to investigate the ethical implications of the digital world on people's lives from a holistic and insightful perspective.

In addition, the perspectives of different actors - such as researchers, digital technology developers, the public, and policy makers, etc. - is not at the core of this thesis. This thesis focuses, instead, on the ethical implications for RECs and other governance approaches to health apps. This is a limitation, as the majority of research with big data is conducted outside of academia and hospitals, and digital health applications are easily accessible to the public as commercial products. Although it is true that each actor plays a fundamental role in the dynamics of ethical considerations and governance, the literature has already explored the perspectives of players such as researchers and developers on this issue (35). In contrast, the implications of big data for oversight mechanisms received far less attention (as seen in Chapters 2, 5 and 9). Therefore, this limitation can simultaneously be interpreted as an attempt to fill a gap in the literature, and to complement the perspectives already explored by other authors. Nevertheless, future research should try to integrate and balance these viewpoints while searching for the middle ground among them.

Regarding the methodological limitations of this thesis, selection bias and subjective bias may arise when reviewing and analyzing literature or policy documents (especially in chapters 2,4,5,8, and 9). These biases can be caused by a range of factors. First, the selected literature may not be exhaustive. In fact, we identified and used only literature written in English, Italian, French and German - the languages spoken by the researchers. Despite this limitation, it should be noted that most modern scientific literature is published in English. Furthermore, although the literature search was conducted using only a few databases, bias was minimized by using the major databases for biomedical and scientific research, and by consulting systematic literature review experts to validate the search strategy. Second, our reviews include papers and policy documents which are not homogeneous in terms of format, content, and quality. However, it is possible for

authors to discuss ethical issues independently from the methodological quality of the study they conduct, the format in which they present the information, and the specific content they address. Nonetheless, we tried to prevent this bias by using uniform (inductive and deductive) criteria for the thematic analysis, and by replicating the data extraction and analysis phase with multiple researchers. Finally, even though this research aims to be international in scope, most of the literature and documents analyzed come from OECD countries, specifically from high-income western countries. This is also true of Chapter 7, which presents a conceptual assessment from an array of perspectives, however, none of these perspectives were from an Asian, African, or South American context. This limitation urges future research to complement the viewpoints explored in this thesis with others from different geographical and socioeconomical settings, population subgroups (e.g., minorities), AI-enabled technology uses, and big data applications. Indeed, cultures and populations around the world may not only have a unique experience of big data and digital health technologies, but also prioritize distinct ethical issues.

Linked to this last point is another methodological limitation, the poor generalizability of the conclusions advanced in the case studies (Chapters 3, 4, 6, and 9). The fact that these chapters focus, for example, only on specific types of apps for digital epidemiology (Chapter 9) or travel medicine (Chapter 4), makes their conclusions ungeneralizable to all health apps. Similarly, Chapter 6 analyzes the perspective of only the Swiss RECs, which represent too small a sample size and constrained territoriality to abstract conclusions at the international level. Nevertheless, it should be noted that the Swiss RECs function in an analogous way (broadly speaking) to RECs in other European countries, the US, Canada, and Australia. This similarity, although only partial, can justify a certain degree of generalization. With appropriate caveats, ethical issues identified in certain health apps may also be extended across health apps and among digital health or medical technologies. For instance, Chapter 3 mentions the ethical implications of using AI in digital medicine (e.g., for diagnosing or preventing diseases). The same ethical implications can be found in those health apps that integrate AI systems. Therefore, it is plausible that a similar range of ethical issues could emerge across different contexts and technologies. However, further research is required to bring more generalizable and statistically significant data that is representative of multiple stakeholders' views on the ethical and governance implications of big data in the health sector.

10.3. Added value of this research

This thesis attests to the lack of adequate attention toward ethical considerations of big data in health apart from privacy and confidentiality. Moreover, these ethical considerations are scattered among the international academic debate and are not addressed in any coherent or comprehensive ethical framework. The absence of such ethical guidance can affect the way in which various actors decide how to use big data and AI-enabled technologies, and how effectively oversight mechanisms can exercise their roles (36). Consequently, potential risks at both the individual and societal level may emerge. The recent the COVID-19 pandemic revealed this issue even more clearly (37), as many actors found themselves without any compass to guide their ethically complex decisions (such as sharing health data to promote the public good while protecting individual rights) (38, 39). In this context, technological solutions (e.g., privacy by design) were insufficient to guide ethical choices, especially when these choices concerned the reconciliation or prioritization of conflicting values among stakeholders (40). In light of recent events, there is an even more urgent need for an ethical guidance that is comprehensive, clear, and rooted in shared values (such as solidarity, dignity, transparency, fairness, and accountability). This thesis also calls for a more robust ethical oversight system that is adaptable to technological change and capable of ensuring ethically aligned use of data.

10.3.1. Policy implications

This thesis uncovers gaps in the current regulatory and oversight landscape. In response to these gaps, the following strategies may help policy makers and regulators to recalibrate and streamline the processes that govern ethical data uses in health research and digital health applications.

Strategy 1: Revise and expand existing ethical standards. This could be the preliminary step toward strengthening the role of ethical oversight and ensuring fair, effective, and safe technological development. Ethical requirements should be extended beyond privacy and product development compliance. As illustrated by our findings, concerns of big data and AI-enabled technologies do not exclusively have an impact at the individual level. On the contrary, group-level harms and risks for the society most often emerge in this context. Therefore, oversight mechanisms should refrain from solely using traditional ethical tools (e.g., informed consent) tailored mostly to protect individual interests. Instead, ethical requirements should guide oversight bodies toward a broader impact-assessment of technologies. As such, a revised governance should promote human rights and shared values, such as transparency, fairness, social justice,

accountability, and dignity. In other words, there is a need for a governance aligned with what society (the various stakeholders - including minorities and underrepresented groups) considers an ethically acceptable use of data and AI-enabled technologies. Therefore, public involvement informing ethical guidance is crucial.

Strategy 2: Translate general ethical guidance into actionable practices. General ethical principles and broad recommendations are often vague and broadly interpretable. Thus, to ensure an ethical use of data and trustworthy digital health development, guidance should come in the form of good practices and easily accessible and operationalizable checklists for stakeholders. Arguing for a clear and implementable governance does not mean that it should be set in stone, instead, it should maintain a degree of flexibility to address emerging data uses and new applications in a timely manner. As this thesis suggests, the market of digital health applications evolves at a fast pace, calling for an adaptive governance approach which could address the associated ethical, social, and legal implications.

Strategy 3: Establish minimum international standards for ethically aligned data uses and quality assurance of digital health applications. This thesis indicates that current governance is fragmented, leaving normative gaps, and forcing various actors to self-regulate. This uncertainty increases confusion among oversight mechanisms and results in inconsistent review approaches within and across fields (e.g., big data research conducted by public vs. private actors, or in an international setting). Furthermore, it raises the risk of malicious actors exploiting regulatory weaknesses to their advantage. In fact, self-certification practices do not demonstrate technology safety, nor do they prove the scientific validity of the evidence, let alone guarantee developer and researcher accountability. Even in the case of good intentions, having to navigate through multiple guidelines and regulations can result in lower quality and unreliable technologies. Therefore, regulatory and ethical standards should be harmonized, thus increasing the transparency of oversight mechanisms by relying on a homogeneous set of rules. Transparent oversight mechanisms may also induce greater public trust both in the oversight mechanisms and in the technologies themselves. Moreover, a cohesive system of norms for health data processing could promote more international collaborations and open data sharing strategies.

Strategy 4: Reform the oversight mechanisms themselves, not just the ethical and normative frameworks. This thesis has a particular focus on the oversight mechanisms in health research and health apps. Concerning the former, the regulations defining the scope of RECs should be

expanded, also providing assessment to those research projects that currently fall outside of the purview of ethical review. Eventually, regulations could suggest the inclusion of specialized and independent ethical committees, distinct from RECs, that provide accreditation to research in big data. Concerning the latter, policy strategies should promote transparent criteria to assess the quality of health apps. In addition, regulation should foster the inclusion of the public in the oversight system, not to legitimize the assessment process, but rather to value the perspective of involved stakeholders. This thesis also indicates that the ethical oversight of health apps should test whether the technical and ethical measures in place to protect individuals and collectives are effective. Accordingly, the oversight mechanisms should include ongoing monitoring and assessment that extends beyond the initial approval of the given research or app. For example, in the context of research, risks arising from publishing research findings should be carefully considered. A system that regularly re-assesses technologies could act quickly to minimize potential risks, but also to mitigate their harmful consequences more effectively once they occur.

10.3.2. Ethical Toolkit

Alongside the above-mentioned policy recommendations, this thesis presents an ethical toolkit that could help fill the gaps left by high-level recommendations and international governance. The toolkit is the final and concrete output of this research; it is informed by research findings and directs stakeholders toward an ethical use of big data, within and outside the field of health, and potentially within and outside of academia.

The toolkit consists of two parts. The first one is a *self-assessment tool* (Figure 10.1). This tool may help RECs, both in Switzerland and abroad, to assess their weaknesses when reviewing big data projects. By using the flowchart, the RECs can recognize and understand their shortcomings. Furthermore, these shortcomings can then be addressed by following the good practices suggested in the *suggested actions* boxes. This tool also sheds light on the types of big data projects that often fall outside the RECs' purview. These are projects relying on anonymized data, happening outside academia, or involving non-traditional actors (such as private companies or lay citizens), and cutting across disciplines, methodologies, geographical locations, and jurisdictions. By mentioning them, the tool draws RECs' attention to regulatory areas in research that may be reviewed and expanded.

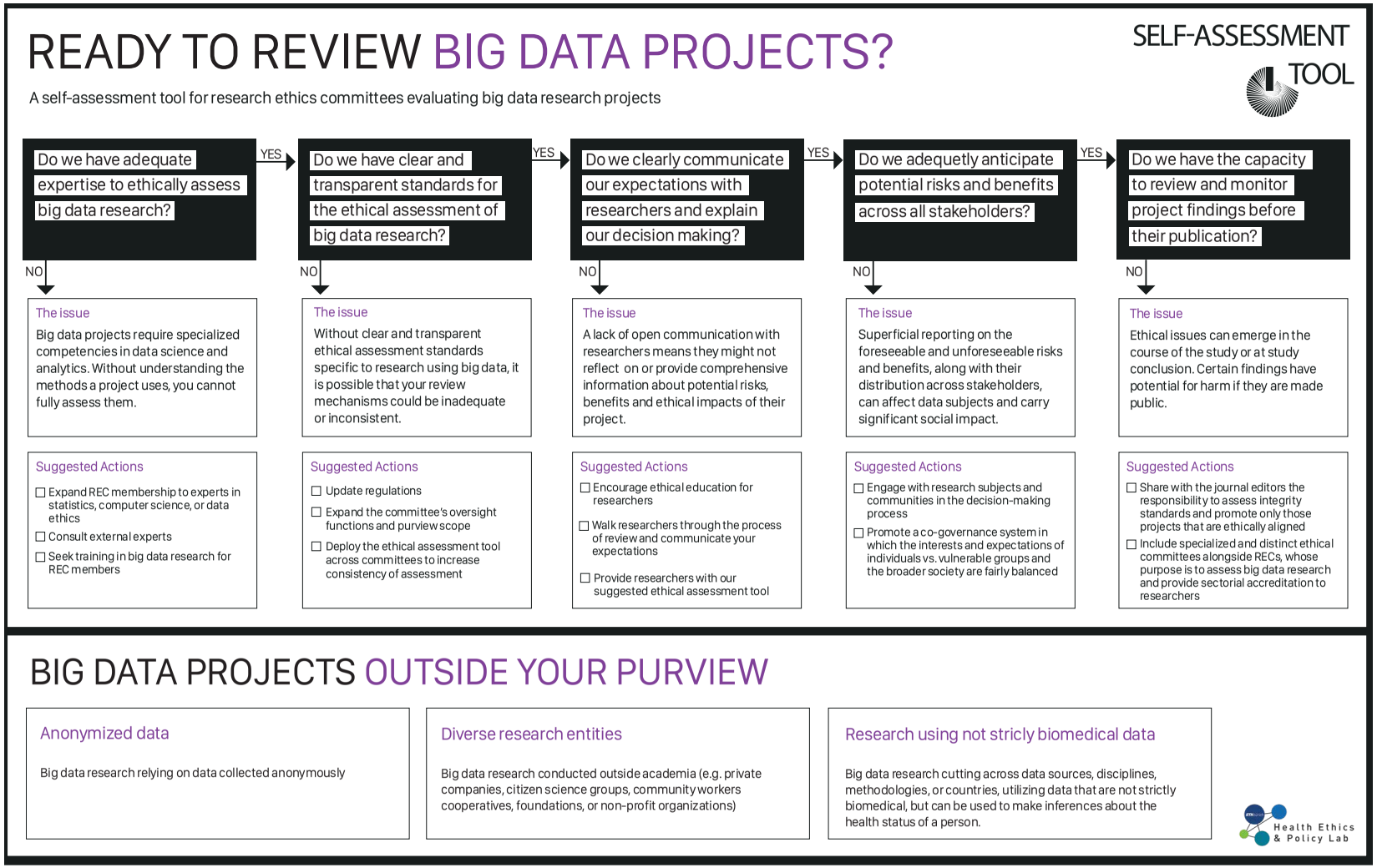
The *ethical assessment tool* (Figure 10.2) forms the second part of the toolkit. This tool is in the form of a checklist which presents relevant categories for conducting ethically aligned research

for stakeholders working with big data. This tool can be used to guide researchers when writing their big data projects' protocols. Researchers, especially those less familiar with ethical considerations, may become more conscious of ethical issues by following a list of questions that orients their assessment. RECs, too, may use this tool for reviewing big data projects by going through the questions and verifying that they are all appropriately addressed in the research project protocol. Hypothetically, RECs and researchers would use the same ethical framework when reviewing and planning big data research, respectively. This could increase the transparency of the review process, while clarifying RECs' expectations and the ethical standards that researchers should meet.

Because of its ease of use, the *ethical assessment tool* may even be deployed outside the academic environment. Developers and researchers in the private sector could find it interesting and convenient to use, while corporations might benefit from clear big data ethics standards when collaborating with universities, especially when submitting collaborative projects for ethics review. Furthermore, private actors could use the checklist to anticipate and address potential big data ethical issues that may arise from their practices. This has the potential to protect companies from reputational harms caused by overlooking ethical issues and their negative consequences. Respecting ethical guidance and promoting ethically aligned decisions beyond privacy protection could potentially make private actors more trustworthy in the eyes of the public.

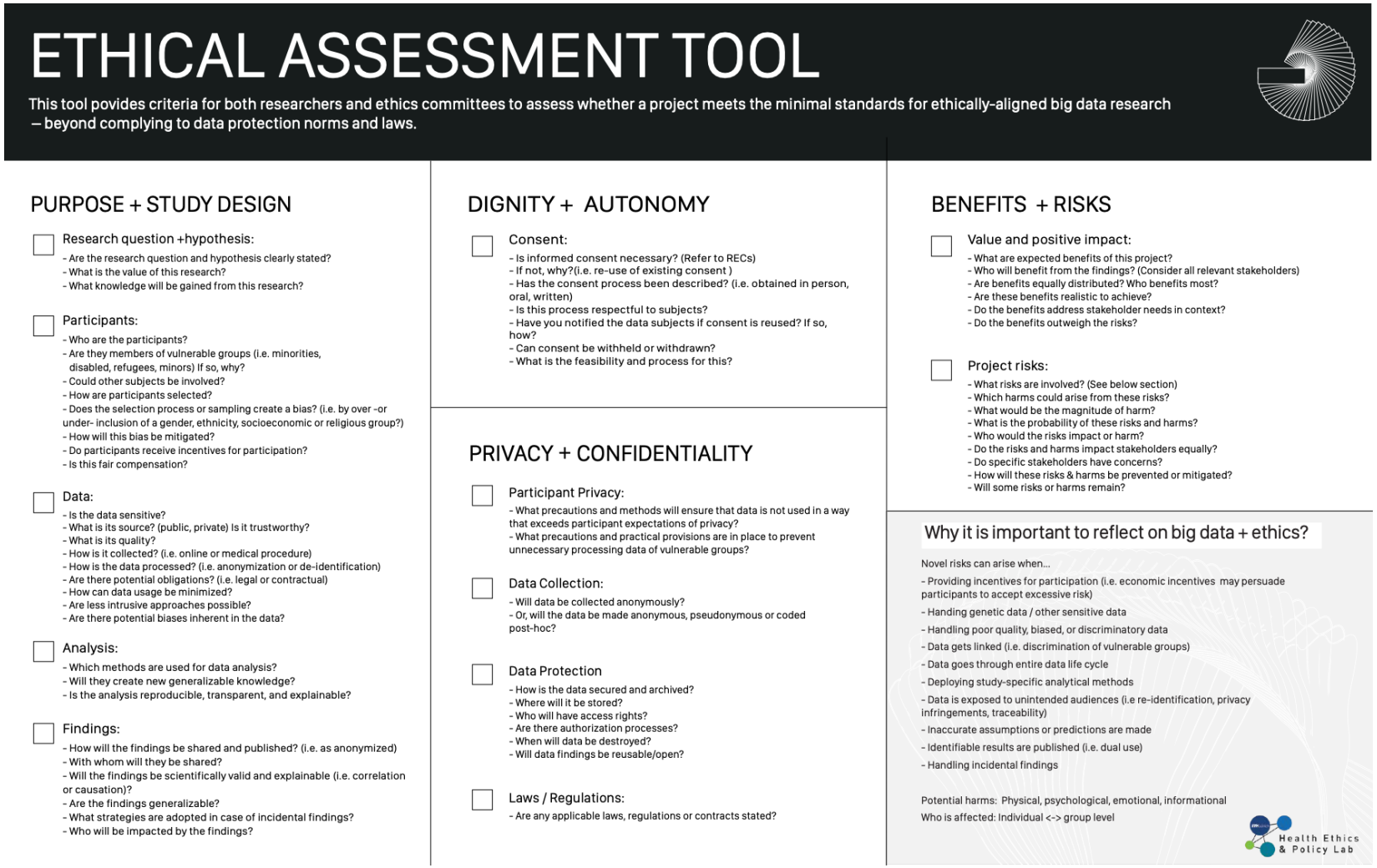
Only a few sporadic examples exist of ethics questionnaires for researchers or RECs focusing on big data uses (41, 42). However, our toolkit is the first to simultaneously address the needs of RECs and researchers. It helps RECs to critically assess their preparedness to face new big data challenges, while simultaneously offering researchers – inside and outside academia – a concrete guide on how to use big data ethically. By doing so, this toolkit supports research and technology development in the big data environment and fills in the grey areas of data governance. We designed this toolkit to be for immediate application. Its guidance can be implemented instantly unlike hard laws which take time to be developed and applied. In doing so, this tool could play a crucial role ensuring that big data practices and research meet high ethical standards.

Figure 10.1: Self-Assessment tool for RECs*



* This tool results from the collaboration of Ferretti, A., Ienca, M., Hurst., S., Sleigh, J., and Vayena, E.

Figure 10.2: Ethical assessment tool for RECs, researchers, and developers*



* This tool results from the collaboration of Ferretti, A., Ienca, M., Hurst., S., Sleigh, J., and Vayena, E.

10.4. Concluding Remarks

Big data and AI-enabled technologies hold enormous potential to revolutionize and improve the health sector. However, to unlock this potential, emerging ethical implications should be urgently addressed. Big data cannot solve this challenge, but we – as a society – have the moral duty to tackle it. To succeed in this quest, we need moral imagination that goes beyond what is prescribed by data privacy regulation and digital health compliance. We must reflect on which vision of future society is preferable and what collective values we want to promote. Consequently, we should assess whether current data usage and technology development are aligned or incompatible with these values. To steer the wheel towards what ought to be done, ethical guidance and oversight mechanisms should be reformed. The evidence uncovered in this work can aid the reform process. A more robust and effective ethical framework can orient researchers, developers, patients, physicians, decision makers, and the public toward ethical choices when using big data in health research and digital health applications.

References

1. Sasubilli SM, Kumar A, Dutt V, editors. Improving Health Care by Help of Internet of Things and Bigdata Analytics and Cloud Computing. 2020 International Conference on Advances in Computing and Communication Engineering (ICACCE); 2020: IEEE.
2. Toh TS, Dondelinger F, Wang D. Looking beyond the hype: Applied AI and machine learning in translational medicine. EBioMedicine. 2019;47:607-15.
3. Miotto R, Wang F, Wang S, Jiang X, Dudley JT. Deep learning for healthcare: review, opportunities and challenges. Briefings in bioinformatics. 2018;19(6):1236-46.
4. Fröhlich H, Balling R, Beerenwinkel N, Kohlbacher O, Kumar S, Lengauer T, et al. From hype to reality: data science enabling personalized medicine. BMC medicine. 2018;16(1):1-15.
5. Obermeyer Z, Emanuel EJ. Predicting the future—big data, machine learning, and clinical medicine. The New England journal of medicine. 2016;375(13):1216.
6. The Atlantic. AI-Driven Dermatology Could Leave Dark-Skinned Patients Behind 2018 [
7. Redwood S, Gill PS. Under-representation of minority ethnic groups in research—call for action. British Journal of General Practice; 2013.
8. Gay V, Leijdekkers P. Bringing Health and Fitness Data Together for Connected Health Care: Mobile Apps as Enablers of Interoperability. J Med Internet Res. 2015;17(11):e260.
9. Abdolkhani R, Gray K, Borda A, DeSouza R. Patient-generated health data management and quality challenges in remote patient monitoring. JAMIA open. 2019;2(4):471-8.
10. Selbst A, Powles J. “Meaningful Information” and the Right to Explanation. In: Sorelle AF, Christo W, editors. Proceedings of the 1st Conference on Fairness, Accountability and Transparency; Proceedings of Machine Learning Research: PMLR; 2018. p. 48--.
11. Wilkinson J, Arnold KF, Murray EJ, van Smeden M, Carr K, Sippy R, et al. Time to reality check the promises of machine learning-powered precision medicine. The Lancet Digital Health. 2020.
12. Anderson M, Anderson SL. How should AI be developed, validated, and implemented in patient care? AMA journal of ethics. 2019;21(2):125-30.
13. Blasimme A, Vayena E. The ethics of AI in biomedical research, patient care and public health. Patient Care and Public Health (April 9, 2019) Oxford Handbook of Ethics of Artificial Intelligence, Forthcoming. 2019.
14. Schiff D, Borenstein J. How should clinicians communicate with patients about the roles of artificially intelligent team members? AMA journal of ethics. 2019;21(2):138-45.
15. Price WN, Gerke S, Cohen IG. How Much Can Potential Jurors Tell Us About Liability for Medical Artificial Intelligence? Journal of Nuclear Medicine. 2021;62(1):15-6.
16. Vayena E, Blasimme A, Cohen IG. Machine learning in medicine: addressing ethical challenges. PLoS medicine. 2018;15(11):e1002689.
17. Health Ethics and policy Lab - ETH Zurich. How the General Data Protection Regulation changes the rules for scientific research European Parliament -Think Thank; 2019 [Available from: https://www.europarl.europa.eu/thinktank/en/document.html?reference=EPRS_STU(2019)634447.
18. Dove ES, Townend D, Meslin EM, Bobrow M, Littler K, Nicol D, et al. Ethics review for international data-intensive research. Science. 2016;351(6280):1399-400.
19. Cavoukian A, Chibba M. Start with privacy by design in all big data applications. Guide to big data applications: Springer; 2018. p. 29-48.

20. Häikiö J, Yli-Kauhaluoma S, Pikkarainen M, Iivari M, Koivumäki T. Expectations to data: Perspectives of service providers and users of future health and wellness services. Health and Technology. 2020:1-16.
21. Roberts JL, Hawkins J. When health tech companies change their terms of service. Science. 2020;367(6479):745-6.
22. Vayena E, Mastroianni A, Kahn J. Caught in the web: informed consent for online health research. Sci Transl Med. 2013;5(173):173fs6.
23. Vayena E, Tasioulas J. The dynamics of big data and human rights: the case of scientific research. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences. 2016;374(2083):20160129.
24. Samuel G, Chubb J, Derrick G. Boundaries Between Research Ethics and Ethical Research Use in Artificial Intelligence Health Research. Journal of Empirical Research on Human Research Ethics. 2021:15562646211002744.
25. Lynch HF, Nicholls S, Meyer MN, Taylor HA, Oversight CtAERE. Of parachutes and participant protection: Moving beyond quality to advance effective research ethics oversight. Journal of Empirical Research on Human Research Ethics. 2019;14(3):190-6.
26. Samuel G, Derrick G, van Leeuwen TN. Ethical challenges around the use of social media data: Views of researchers and research ethics committee members. 2018.
27. Torous J, Nebeker C. Navigating ethics in the digital age: introducing connected and open research ethics (CORE), a tool for researchers and institutional review boards. Journal of medical Internet research. 2017;19(2):e38.
28. Marelli L, Lievevrouw E, Van Hoyweghen I. Fit for purpose? The GDPR and the governance of European digital health. Policy studies. 2020;41(5):447-67.
29. Public Health England. Criteria for health app assessment 2021 [Available from: https://www.gov.uk/government/publications/health-app-assessment-criteria/criteria-for-health-app-assessment.
30. Magrabi F, Habli I, Sujan M, Wong D, Thimbleby H, Baker M, et al. Why is it so difficult to govern mobile apps in healthcare? BMJ Health and Care Informatics. 2019.
31. Bates DW, Landman A, Levine DM. Health apps and health policy: what is needed? Jama. 2018;320(19):1975-6.
32. Gordon WJ, Landman A, Zhang H, Bates DW. Beyond validation: getting health apps into clinical practice. npj Digital Medicine. 2020;3(1):14.
33. Blasimme A. Producing health: Georges Canguilhem and the contemporary clinical discourse about aging. The Normal and the Pathological: Essays on Georges Canguilhem: Hermann; 2020.
34. Ienca M, Ignatiadis K. Artificial intelligence in clinical neuroscience: methodological and ethical challenges. AJOB neuroscience. 2020;11(2):77-87.
35. Favaretto M, De Clercq E, Schneble CO, Elger BS. What is your definition of Big Data? Researchers' understanding of the phenomenon of the decade. PloS one. 2020;15(2):e0228987.
36. Reddy S, Allan S, Coghlan S, Cooper P. A governance model for the application of AI in health care. Journal of the American Medical Informatics Association. 2020;27(3):491-7.
37. Gardner L, Ratcliff J, Dong E, Katz A. A need for open public data standards and sharing in light of COVID-19. The Lancet Infectious Diseases. 2020.
38. Vayena E. Value from health data: European opportunity to catalyse progress in digital health. The Lancet. 2021;397(10275):652-3.
39. Ienca M, Vayena E. On the responsible use of digital data to tackle the COVID-19 pandemic. Nature medicine. 2020;26(4):463-4.
40. Ngan OM, Kelmenson AM. Using Big Data Tools to Analyze Digital Footprint in the COVID-19 Pandemic: Some Public Health Ethics Considerations. Asia Pacific Journal of Public Health. 2020:1010539520984360.

41. Future of Privacy Forum. Beyond IRBs: Designing Ethical Review Processes for Big Data Research 2015 [Available from: https://fpf.org/wp-content/uploads/2017/01/Beyond-IRBs-Conference-Proceedings_12-20-16.pdf.
42. University of Leicester. Ethics checklist 2016 [Available from: https://www2.le.ac.uk/institution/ethics/resources/checklist.

List of Abbreviations

AAL:	Ambient Assisted Living
AI:	Artificial Intelligence
API:	Application Programming Interface
CANDALS:	Citizenship, Ability, Neurotypicality/Neurodiversity, Disability, Age, Literacy and/or fluency, and Size, BMI, or body habitus
CEC:	Central Ethic Committee
CENTRAL:	Cochrane Central Register of Controlled Trials
CoC:	Code of Conduct
DPA:	Data Protection Authority
DPD:	Data Protection Directive
DP-3T:	Decentralized Privacy-Preserving Proximity Tracing
DTC:	Digital Contact Tracing
EC:	European Commission
ELSI:	Ethical Legal and Social Implications
ERB(s):	Ethics Review Board(s)
ERC(s):	Ethics Review Committee(s)
FDA:	Food and Drug Administration
GAEN:	Google/Apple Exposure Notification system
GDPR:	General Data Protection Regulation
GPS:	Global Positioning System
HAS:	Haute Autorité de Santé
HIPAA:	Health Insurance Portability and Accountability Act
HIT:	Health Information Technologies
HRA:	Human Research Act
ICDPPC:	International Conference on Data Protection and Privacy Commissioners
IP:	Internet Protocol
IRB(s):	Institutional Review Board(s)
ITIT:	Illness Tracking in Travelers
LMICs:	Low- and Middle-Income Countries
ML:	Machine Learning
mERA:	Mobile Evidence Reporting and Assessment

mHEALTH:	Mobile Health
NGOs:	Non-governmental organization
NHS:	National Health Service
NIH:	National Institute of Health
OECD:	Organisation for Economic Co-operation and Development
PEPP-PT:	Pan-European Privacy-Preserving Proximity Tracing
REC(s):	Research Ethics Committee(s)
RCT:	Research Clinical Trial
SAMS:	Swiss Academy of Medical Sciences
SANRA:	Scale for the Assessment of Narrative Review Articles
STROBE:	Strengthening the Reporting of Observational studies in Epidemiology
UNESCO:	United Nations Educational, Scientific and Cultural Organization
UK:	United Kingdom
US:	United States
VFR:	Visiting Friends and Relatives
WHO:	World Health Organization

List of Tables

List of Figures

Appendixes

Appendix 1 Chapter 4: Search Strings By Database

Medline

[exp Travel Medicine/ or ((travel* or tourist* or trip or trips) and (health or medicine or mhealth or illness* or infection* or symptom* or ehealth or epidemiolog* or disease*)).mp.] AND [exp Mobile Applications/ or ((mobile adj2 (app or apps or application* or phone* or device*)) or tracking or technolog* or smartphone*).mp.] AND [exp Ethics/ or (value* or bioethic* or ethic* or consent* or "e-consent*" or autonom* or empower* or privacy or confidential* or "data protection" or anonym* or "de-identification" or securit* or proportionalit* or trust* or transparen* or fairness or justice* or equit* or (public adj2 (good* or benefit*)) or solidarity or "non-maleficence*" or accountab* or legitimation* or governance* or oversight*).mp.] limit to (english or german or french or italian)

Embase

['emporiatrics'/exp OR ((travel*:ti,ab,kw OR tourist*:ti,ab,kw OR trip:ti,ab,kw OR trips:ti,ab,kw) AND(health:ti,ab,kw OR medicine:ti,ab,kwOR mhealth:ti,ab,kw OR illness*:ti,ab,kw OR infection*:ti,ab,kw OR symptom*:ti,ab,kw OR ehealth:ti,ab,kw OR epidemiolog*:ti,ab,kw OR disease*:ti,ab,kw))] AND ['mobile application'/exp OR ((mobile NEAR/2 (app OR apps OR application* OR phone* OR device*)):ti,ab,kw) OR tracking:ti,ab,kw OR technolog*:ti,ab,kw OR smartphone*:ti,ab,kw] AND ['ethics'/exp OR value*:ti,ab,kw OR bioethic*:ti,ab,kw OR ethic*:ti,ab,kw OR consent*:ti,ab,kw OR 'e-consent*':ti,ab,kw OR autonom*:ti,ab,kw OR empower*:ti,ab,kw OR privacy:ti,ab,kw OR confidential*:ti,ab,kw OR 'data protection':ti,ab,kw OR anonym*:ti,ab,kw OR 'de- identification':ti,ab,kw OR securit*:ti,ab,kw OR proportionalit*:ti,ab,kw OR trust*:ti,ab,kw OR transparen*:ti,ab,kw OR fairness:ti,ab,kw OR justice*:ti,ab,kwOR equit*:ti,ab,kw OR ((public NEAR/2 (good* OR benefit*)):ti,ab,kw) OR solidarity:ti,ab,kw OR 'non-maleficence*':ti,ab,kw OR accountab*:ti,ab,kw OR legitimation*:ti,ab,kw OR governance*:ti,ab,kw OR oversight*:ti,ab,kw] NOT [conference abstract]/lim AND ([english]/lim OR [german]/lim OR [french]/lim OR [italian]/lim)

Scopus

(TITLE-ABS-KEY ((travel* OR tourist* OR trip OR trips) AND (health OR medicine OR mhealth OR illness* OR infection* OR symptom* OR ehealth OR epidemiolog* OR disease*))) AND (TITLE-ABS-KEY (mobile W/2 (app OR apps OR application* OR phone* OR device*) OR tracking Or technolog* OR smartphone)) AND (TITLE-ABS-KEY (value* OR bioethic* OR ethic* OR consent* OR OR “e-consent*” OR autonom* OR empower* OR privacy OR confidential* OR “data protection” OR anonym* OR “de-identification” OR securit* OR proportionalit* OR trust* OR transparen* OR fairness OR justice* OR equit* OR (public W/2 (good* OR benefit*)) OR solidarity OR “non-maleficence*” OR accoutnab* OR legitimation* OR governance* OR oversight*)) AND (LIMIT-TO (LANGUAGE , “English”) OR LIMIT-TO (LANGUAGE , “French”))

Web of Science

(TS=((travel* OR tourist* OR trip OR trips) AND (health OR medicine OR mhealth OR illness* OR infection* OR symptom* OR ehealth OR epidemiolog* OR disease*))) AND (TS=(mobile NEAR/2 (app OR apps OR application* OR phone* OR device*) OR tracking OR technolog* OR smartphone)) AND TS=(value* OR bioethic* OR ethic* OR consent* OR "e-consent*" OR autonom* OR empower* OR privacy OR confidential* OR "data protection" OR anonym* OR "de-identification" OR securit* OR proportionalit* OR trust* OR transparen* OR fairness OR justice* OR equit* OR (public NEAR/2 (good* OR benefit*)) OR solidarity OR "non-maleficence*" OR accountab* OR legitimation* OR governance* OR oversight*) AND LANGUAGE: (English OR French OR German OR Italian) Indexes=SCI-EXPANDED,SSCI,A&HCI,CPCI-S,CPCI-SSH,BKCI-S,BKCI-SSH,ESCI,CCR-EXPANDED,IC Timespan=All years

Cochrane

[((travel* OR tourist* OR trip OR trips) AND (health OR medicine OR mhealth OR illness* OR infection* OR symptom* OR ehealth OR epidemiolog* OR disease*)):ti,ab,kw] AND [(mobile NEAR/2 (app OR apps OR application* OR phone* OR device*) OR tracking OR technolog* OR smartphone*):ti,ab,kw] AND [(value* OR bioethic* OR ethic* OR consent* OR "e-consent*" OR autonom* OR empower* OR privacy OR confidential* OR "data protection" OR anonym* OR "de-identification" OR securit* OR proportionalit* OR trust* OR transparen* OR fairness OR justice* OR equit* OR (public NEAR/2 (good* OR benefit*)) OR solidarity OR "non-maleficence*" OR accountab* OR legitimation* OR governance* OR oversight*):ti,ab,kw]

IEEE Xplore

(((travel* OR tourist*) AND (health OR disease* OR ill*) AND (mobile OR application OR smartphone*) AND (value* OR ethic* OR consent OR privacy)))

Science Direct

((travel OR tourist) AND (smartphone OR application) AND (health OR illness) AND (value OR ethics OR ethical))

medRxiv

Subject area: Medical Ethics
Title: Travel tourist (any)
Abstract or Title: smartphone application mobile (any)

SSRN

travel smartphone health in title, abstract, keywords

Appendix 2 Chapter 4: Quality assessment of included papers

Author	**Year**	**Study Type**	**Rating system**	Title and Abstract (max 1)	Intro (max 2)	Methods (max 9)	Results (max 5)	Discussion (max 4)	Other (max 1)	Total (max 22)	Point 1 (max 2)	Point 2 (max 2)	Point 3 (max 2)	Point 4 (max 2)	Point 5 (max 2)	Point 6 (max 2)	Total (max 12)
Baroutsou	2020	Cohort Study	STROBE	1	2	8	5	4	1	21							0
Du	2020	Qualitative Analysis	SANRA							0	2	2	0	2	2	2	10
Famham	2018	Cohort Study	STROBE	1	2	8	5	3	1	20							0
Lai	2019	Qualitative Analysis	SANRA							0	2	2	2	2	1	2	11
Sethia	2018	Qualitative Analysis	SANRA							0	2	2	0	2	2	2	10
Subramaniyaswamy	2018	Qualitative Analysis	SANRA							0	1	2	0	2	2	2	9
Seed	2016	Qualitative Analysis	SANRA								2	1	2	2	0	2	9

Appendix 1 Chapter 5: List of included documents

Title	Issued by	Year	Link	Country of issuer	Type of issuer	Targeted stakeholder group	Recommenda-tions for IRBs	Prominent health focus
Universal principles of data ethics. 12 guidelines for developing ethics codes	Accenture	2016	https://www.accenture.com/t20160629T012639Z__w__/us-en/_acnmedia/PDF-24/Accenture-Universal-Principles-Data-Ethics.pdf	Ireland	Private Sector	Multiple stakeholders	YES	NO
Guiding Principles for the Ethical Use of Data	Acxiom Corporation	2019	https://marketing.acxiom.com/rs/982-LRE-196/images/ACX001_EthicalUseofData.pdf	USA	Private Sector	Multiple stakeholders	NO	NO
National and Transnational Security Implications of Big Data in the Life Sciences	American Association for the Advancement of Science (AAAS)	2014	https://www.aaas.org/sites/default/files/AAAS-FBI-UNICRI_Big_Data_Report_111014.pdf	USA	Professional Organisation	Government regulators	NO	YES
Ethical Guidelines for Statistical Practice	American Statistical Association	2016	http://www.amstat.org/asa/files/pdfs/EthicalGuidelines.pdf	USA	Professional Organisation	Professionals	NO	YES
ACM Code of Ethics and Professional Conduct	Association for Computing Machinery (ACM)	2018	https://www.acm.org/about-acm/acm-code-of-ethics-and-professional-conduct	International	Professional Organisation	Professionals	NO	NO
Ethical decision-making and Internet research 2.0: Recommendations from the AoIR ethics working committee	Association of Internet Researchers, Ethics Working Committee	2012	http://aoir.org/ethics/	International	Professional Organisation	Government regulators	YES	NO
Big Data Roadmap	Association of the British Pharmaceutical Industry	2013	http://www.abpi.org.uk/publications/big-data-road-map	UK	Professional Organisation	Private companies	NO	YES
Guide to big data and the Australian Privacy Principles	Australian Government, Office of the Australian Information Commissioner	2016	https://www.oaic.gov.au/engage-with-us/consultations/guide-to-big-data-and-the-australian-privacy-principles/consultation-draft-guide-to-big-data-and-the-australian-privacy-principles	Australia	National Government Institution	Multiple stakeholders	NO	YES
Code of Good Practice	British Computer Science Association	2011	http://www.bcs.org/upload/pdf/cop.pdf	UK	Professional Organisation	Professionals	NO	NO
Big data roadmap and cross-disciplinary community for addressing societal Externalities	BYTE Project	2014	http://new.byte-project.eu/wp-content/uploads/2014/02/D7.3-Final-report-FINAL.pdf	EU	Intergovernmental and Supra-governmental Institutions	Multiple stakeholders	NO	NO
Big Data Analytics in Health (White Paper)	Canada Health Infoway	2013	https://www.infoway-inforoute.ca/en/component/edocman/1246-big-data-analytics-in-health-white-paper-full-report/view-document?Itemid=0	Canada	NGO	Multiple stakeholders	NO	YES
Better Information for Improved Health: A Vision for Health System Use of Data in Canada	Canadian Institute for Health Information	2013	https://www.cihi.ca/en/hsu_vision_report_en.pdf	Canada	National Government Institution	Multiple stakeholders	NO	YES
Health Big Data in the Clinical Context	Centre for Democracy and Technology	2015	https://cdt.org/files/2015/04/Health-Big-Data-in-the-Clinical-Context.pdf	USA	NGO	Private companies	NO	YES
Big Data and Analytics: Seeking Foundations for Effective Privacy Guidance	Centre for Information Policy Leadership (CIPL)	2013	https://www.hunton.com/files/Uploads/Documents/News_files/Big_Data_and_Analytics_February_2013.pdf	International	Think Tank Platform	Private companies	NO	YES
Big Data Security and Privacy Handbook: 100 best practices in big data security and privacy	Cloud Security Alliance	2016	https://downloads.cloudsecurityalliance.org/assets/research/big-data/BigData_Security_and_Privacy_Handbook.pdf	International	NGO	Private companies	NO	NO
Tecnologie dell'informazione e della	CNB (comitato nazionale per la bioetica)	2016	http://www.quotidianosanita.it/allegati/allegato7545104.pdf	Italy	National Government Institution	Government regulators	NO	YES

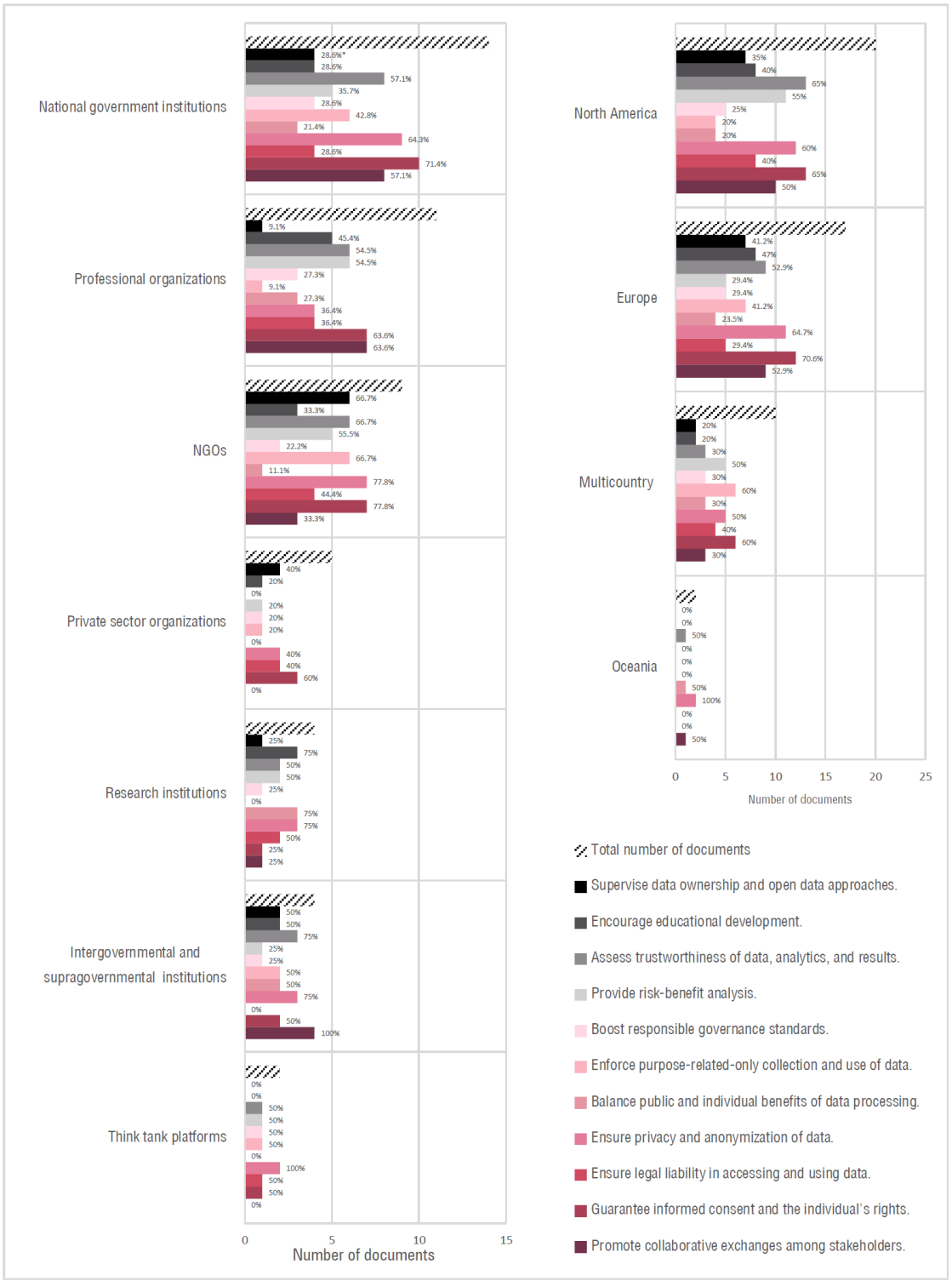
comunicazione: profili bioetici.								
Perspectives on Big Data, Ethics, and Society	Council for Big Data, Ethics and Society	2016	https://bdes.datasociety.net/wp-content/uploads/2016/05/Perspectives-on-Big-Data.pdf	USA	Research Institution	Government regulators	YES	NO
Guidelines on the protection of individuals with regard to the processing of personal data in a world of Big Data	Council of Europe	2017	https://rm.coe.int/16806ebe7a","https://rm.coe.int/16806ebe7a	EU	Intergovernmental and Supra-governmental Institutions	Multiple stakeholders	YES	NO
Supporting Ethical Data Research: An Exploratory Study of Emerging Issues in Big Data and Technical Research	Data & Society	2016	https://www.datasociety.net/pubs/sedr/SupportingEthicsDataResearch_Sept2016.pdf	USA	Research Institution	Professionals (researchers)	YES	NO
Data Science Code of Conduct	Data Science Association	Accessed April 2019	http://www.datascienceassn.org/code-of-conduct.html	USA	Professional Organisation	Professionals	NO	NO
The Menlo Report: Ethical Principles Guiding Information and Communication Technology Research	Department of Homeland Security (DHS) - Center for Applied Internet Data Analysis	2012	https://www.caida.org/publications/papers/2012/menlo_report_actual_formatted/menlo_report_actual_formatted.pdf	USA	National Government Institution	Multiple stakeholders	YES	NO
Guidance for Incorporating Big Data into Humanitarian Operations	Digital Humanitarian Network	2015	http://digitalhumanitarians.com/sites/default/files/resource-field_media/IncorporatingBigDataintoHumanitarianOps-2015.pdf	International	NGO	NGO members	NO	NO
The Use of Big Data in Public Health Policy and Research	European Commission	2014	https://ec.europa.eu/health/sites/health/files/ehealth/docs/ev_20141118_co07b_en.pdf	EU	Intergovernmental and Supra-governmental Institutions	Multiple stakeholders	NO	YES
Big data in Healthcare – what role for the EU?	European Health Parliament	2017	https://www.healthparliament.eu/wp-content/uploads/2017/10/Big-data-in-Healthcare-the-experience-and-results-from-the-European-Health-Parliament.pdf	EU	Intergovernmental and Supra-governmental Institutions	Government regulators	NO	YES
Big Data: A tool for inclusion or exclusion? (FTC Report)	Federal Trade Commission	2016	https://www.ftc.gov/system/files/documents/reports/big-data-tool-inclusion-or-exclusion-understanding-issues/160106big-data-rpt.pdf	USA	National Government Institution	Private companies	NO	NO
Benefit-Risk Analysis for Big Data Projects	Future of Privacy Forum	2014	https://fpf.org/wp-content/uploads/FPF_DataBenefitAnalysis_FINAL.pdf	USA	Think Tank Platform	Multiple stakeholders	YES	YES
Health Big Data Recommendations	HITPC (Health IT Policy Committee), Health Information Technology Advisory Committee (HITAC)	2015	https://www.healthit.gov/sites/default/files/facas/HITPC_Health_Big_Data_Report_FINAL.pdf	USA	National Government Institution	Government regulators	YES	YES
White Paper on "Ethics for big data and analytics"	IBM	2014	http://www.ibmbigdatahub.com/sites/default/files/whitepapers_reports_file/TCG%20Study%20Report%20-%20Ethics%20for%20BD%26A.pdf	USA	Private Sector	Multiple stakeholders	NO	NO
Big data, artificial intelligence, machine learning and data protection	ICO (Information Commissioner's Office)	2017	https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf	UK	National Government Institution	Private companies	YES	YES
IEEE Code of ethics	IEEE	Accessed April 2019	http://www.ieee.org/about/corporate/governance/p7-8.html	International	Professional Organisation	Professionals	NO	NO
Code of Ethics/Conduct	INFORM for the Certified Analytics Professional	Accessed April 2019	https://www.certifiedanalytics.org/ethics.php	USA	Professional Organisation	Professionals	NO	NO
Big Data Guidelines	Information and Privacy Commissioner of Ontario	2017	https://www.ipc.on.ca/wp-content/uploads/2017/05/bigdata-guidelines.pdf	Canada	National Government Institution	Government regulators	YES	NO

Building Ethics into Privacy Frameworks for Big Data and AI	International Association of Privacy Professionals	2018	https://iapp.org/media/pdf/resource_center/BUILDING-ETHICS-INTO-PRIVACY-FRAMEWORKS-FOR-BIG-DATA-AND-AI-UN-Global-Pulse-IAPP.pdf	International	Professional Organisation	Multiple stakeholders	YES	NO
Location Data Privacy: Guidelines, Assessment & Evaluations	Location Forum	2013	https://iapp.org/media/pdf/resource_center/LocationDataPrivacyGuidelines_v2.pdf	International	NGO	Multiple stakeholders	NO	YES
Leitlinien für den Big-Data-Einsatz im Überblick Chancen und Verantwortung	Nationaler IT-Gipfel	2015	https://www.digitale-technologien.de/DT/Redaktion/DE/Downloads/Publikation/Smart_Data_Positionspapier_BigData_Leitlinien.pdf?__blob=publicationFile&v=7	Germany	National Government Institution	Multiple stakeholders	NO	YES
The big data dilemma	Nuffield Council on Bioethics	2015	http://nuffieldbioethics.org/wp-content/uploads/Big-Data-dilemma-Nuffield-Council-on-Bioethics-September-2015.pdf	UK	NGO	Government regulators	NO	YES
The collection, linking and use of data in biomedical research and health care: ethical issues	Nuffield Council on Bioethics	2015	http://nuffieldbioethics.org/wp-content/uploads/Biological_and_health_data_web.pdf	UK	NGO	Multiple stakeholders	YES	YES
Human Subjects Research Implications of "Big Data" Studies	Office for Human Research Protections	2015	https://www.hhs.gov/ohrp/sachrp-committee/recommendations/2015-april-24-attachment-a/index.html	USA	National Government Institution	Government regulators	YES	YES
An Enterprise Architect's Guide to Big Data	Oracle	2016	http://www.oracle.com/technetwork/topics/entarch/articles/oea-big-data-guide-1522052.pdf	USA	Private Sector	Private companies	NO	YES
The Opportunities and Ethics of Big Data	Royal Statistical Society	2015	http://www.rss.org.uk/Images/PDF/influencing-change/2016/rss-report-opps-and-ethics-of-big-data-feb-2016.pdf	UK	Professional Organisation	Professionals	YES	NO
The Open Data Era in Health and Social Care	The GovLab (for NHS England)	2014	http://images.thegovlab.org/wordpress/wp-content/uploads/2014/10/nhs-full-report-21.pdf	UK	National Government Institution	Multiple stakeholders	NO	YES
Big Data Big Possibilities: How Australia Can Use Big Data for Better Healthcare	THE McKell Institute	2016	https://www.allens.com.au/pubs/pdf/healthcare/Healthcare-McKellReport.pdf	Australia	Research Institution	Government regulators	NO	YES
Big data and data sharing: Ethical issues	UK Data Service	2017	https://www.ukdataservice.ac.uk/media/604711/big-data-and-data-sharing_ethical-issues.pdf	UK	National Government Institution	Professionals	YES	YES
Data Ethics Framework	UK Government, Cabinet Office	2018	https://www.gov.uk/government/publications/data-science-ethical-framework	UK	National Government Institution	Government regulators	NO	NO
Integrating Big Data into the Monitoring and Evaluation of Development Programmes	UN Global Pulse	2016	http://unglobalpulse.org/sites/default/files/IntegratingBigData_intoMEDP_web_UNGP.pdf	International	NGO	NGO members	NO	NO
Draft Report on Big Data and Health	UNESCO IBC	2017	http://unesdoc.unesco.org/images/0024/002487/248724E.pdf	International	NGO	Multiple stakeholders	YES	YES
The federal big data research and development strategic plan	US Subcommittee on Networking and Information Technology Research and Development	2016	https://www.nitrd.gov/PUBS/bigdatardstrategicplan.pdf	USA	National Government Institution	Government regulators	NO	YES
Big Data: Ethische Fragen	Vodafone Institut für Gesellschaft und Kommunikation	2016	http://www.vodafone-institut.de/wp-content/uploads/2016/10/Big-Data_Ethische-Fragen.pdf	Germany	Private Sector	Multiple stakeholders	NO	NO
How should health data be used? Privacy, secondary use, and big data sales	Yale University, Institution of Social and Policy Studies	2014	http://bioethics.yale.edu/sites/default/files/files/ISPS14-025.pdf	USA	Research Institution	Multiple stakeholders	NO	YES

Appendix 2 Chapter 5: Inclusion and exclusion criteria of sources

Sources considered	Types	Documents published online or websites featuring materials such as policy documents, soft-law documents, best-practice guidelines, reports, declarations, technical specifications and standards, technical and commercial documentation, and official recommendation documents
	Issuers	National government or international institutions, the private sector (e.g., companies and corporations), NGOs, nonprofit organizations, academic and research institutions, and professional organizations
	Languages	English, Italian, French, German, and Greek (the languages spoken by the authors)
	Content	Documents referring to the ethics of big data and providing normative recommendations or best practices
Sources excluded	Types	Blog articles, academic articles, letters to the editors, journalistic articles, legislation, books, conference proceedings, dissertations, videos, images, and audio recordings and podcasts
	Issuers	Single authors
	Languages	Others than those mentioned above
	Content	Documents that do not mention any ethical topic related to big data or do not provide best practices and recommendations

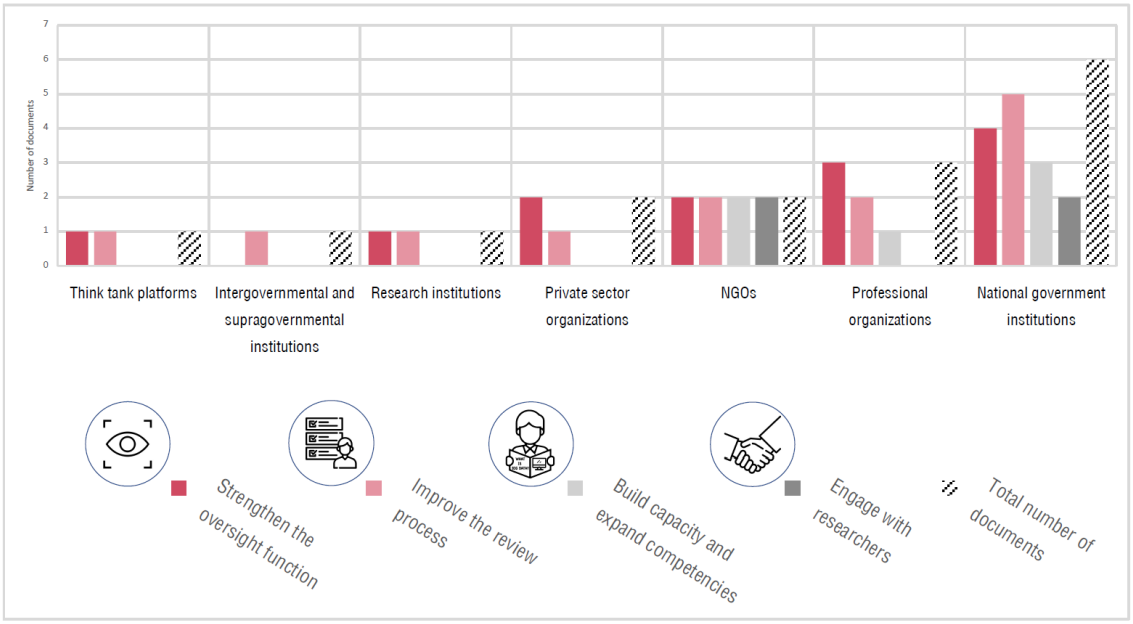
Appendix 3 Chapter 5: Substantive Recommendations for IRBs in Relation to the Type of Issuer and the Issuer's Continent



Each percentage refers to the documents within a specific category (e.g., North America or Europe) that mentioned that specific recommendation (e.g., encourage educational development or provide risk-benefit analysis).

Our analysis identified few peculiarities concerning the substantive ethical recommendations issued (a) in different continents and (b) by various stakeholders. With respect to different continents, almost 60% of North American documents recommended practicing risk- benefit analysis when using big data for research, compared to 30% of European documents. Furthermore, the topic of enforcing purpose-related-only collection and use of data was only partially discussed within European (40%) and North American (20%) documents compared to international documents (60%). The highest proportion of documents including recommendations on ethical educational development for researchers using big data was provided by European documents. When analyzing with respect to stakeholders, intergovernmental and supragovernmental institutions highlighted—more than any other issuer—the importance of collaborative exchange between stakeholders. Our analysis also showed that 80% of research institutions' documents offered recommendations about balancing public and individual benefits of data processing, compared to the 20% of documents issued by national government institutions. Finally, the data breakdown displayed a lack of substantive recommendations by the private sector. For example, the suggestion to assess the trustworthiness of data, the analytical processes, and the results, is discussed by at least 50% of documents released by every other type of stakeholder, but it is not discussed at all in the documents from the private sector.

Appendix 4 Chapter 5: Procedural Recommendations for IRBs in Relation to the Type of Issuer



Among all the stakeholders providing procedural recommendations, only national government institutions and research institutions mentioned all four recommendations. In general, there was high heterogeneity across the analyzed documents. For example, none of the documents issued by professional organizations talked about the necessity for IRBs to engage more with researchers, while documents issued by both research institutions and national governments provided this suggestion. Furthermore, the document issued by the private sector focused exclusively on the necessity to strengthen the oversight function of IRBs, while all the other issuers mentioned at least one other recommendation.

Appendix 6.1: Interview guide

Prior to the interview, a study's investigator will provide an overview of the research purpose and will remind to the participant the confidentiality and anonymity measures adopted in the research. Also, the study investigator will ask to get permission for tape recording.

INTRODUCTION: Respondent's position/function in the Ethical Committee
• What is your professional/scientific background? How often do you serve in this EC?
TOPIC 1: Respondent's understanding of biomedical big data and previous experience in this respect
• What makes you consider a project as "big data"? Are there cases where you are uncertain about whether a project is big data or not? • Does your EC (regularly or occasionally) review big data projects? • If yes, how often? How many of those involved biomedical data? • If no, why? Who does review them instead? • Are online research projects and studies involving publicly available data repositories reviewed by your EC? Should they?
TOPIC 2: Respondent's opinion concerning the promises and challenges brought by biomedical big data
• What are the major scientific benefits that you see associated with biomedical uses of big data? Is there a social need to maximize data availability for research? • What are, in your view, the major ethical and social challenges associated with using big data for research? Any issue specific to healthcare research? (If the participant immediately links the answer to healthcare research, ask whether there are more general issues outside the medical context) • Do you feel that biomedical big data projects pose unprecedented/novel/unique ethical challenges? If yes, which ones? If not, do you think they change existing challenges? • Do you see any particular impact on privacy? Where do you see a fair balance between the social need of maximizing data and the individual need of protecting privacy? • How can individuals consent to the use of their data? How should they? • How do you define "risk" in big-data projects? • How do you assess risk-benefit in healthcare big data projects? • How would define "minimal risk" in relation to biomedical big data research? How do you assess minimal risk in other research context?
TOPIC 3: Existing guidelines and criteria adopted to handle biomedical big data related issues
• Do you follow any specific guidelines to assess biomedical big data projects? If yes, which ones? • Are you aware of guidelines from national or international organizations? • When evaluating a big data project in healthcare, what do you mainly look at? Data type? Data volume? Data collection methods? Analytic methods? • Have you ever heard of algorithmic transparency?
TOPIC 4: Assessing respondents' needs for guidelines in relation to big data

• Do you feel that your EC is adequately equipped to evaluate biomedical big data projects? • If not, what expertise, tool or mechanism would be required? • Do you think the EC has the responsibility to evaluate big data projects? If yes, explain why. If not, which authority should do that instead? (Prompt: data protection office?) • Do you think novel review bodies are needed? If so, do you see them as complementary or substitutive of EC?
TOPIC 5: Respondent's suggestions to develop an inclusive guideline policy concerning big data in healthcare
• If you could contribute to the drafting of new guidelines, what would your main recommendations be? • Which values would be paramount? • Who do you think should develop such guidelines (e.g. WHO, national govs, private corporations etc.)? At which level (international vs national)?

Appendix 1 Chapter 8: Sources and Method

Sources

Nine OECD countries were selected as case studies: Australia, Canada, France, Germany, Italy, New Zealand, Switzerland, the United Kingdom (UK), and the United States (US). Data Protection Authorities (DPAs) in these countries had all reported guidance relating to mobile apps in response to the 2017 Census of the International Conference on Data Protection and Privacy Commissioners (ICDPPC). The ICDPPC Census had received a positive reply from DPAs in 13 OECD countries (Australia, Canada, Germany, Netherlands, Switzerland, Slovenia, Italy, Sweden, United Kingdom, United States, France, New Zealand, and Norway). We contacted all DPAs in these countries and received confirmation of specific guidance from the nine countries included in this study and from the European Commission (EC). All of the guidelines were issued by national DPAs and published between 2012 and 2019, with the exception of the German guideline which was developed by the local DPA of the Bayer region and the Swiss one developed by the canton of Zurich. The Italian DPAs had primarily developed an online brief to assist app users (Italy's "vademecum for consumers") and was therefore not included in the extended analysis. The other missing DPAs, when contacted, could not confirm specific national guidance (Slovenia) and referred back to European Commission Guidance (Netherlands, Norway, and Sweden).

In addition, we searched for guidance on health apps issued by national health authorities in these same countries. We excluded from our search professional organisations and other non-state authorities. We found two documents. The first, produced by health authorities in France by the High Authority of Health (HAS), the second by the UK National Health Services (NHS). We also searched for guidance issued by the European Commission (EC) and the World Health Organisation (WHO). We identified and included in the analysis two documents: the "Privacy Code of Conduct on mobile health apps" (CoC) developed by the EC and the "Guidelines for reporting of health interventions using mobile phones: mobile health (mHealth) evidence reporting and assessment (mERA) checklist" of the WHO mHealth Technical Evidence Review Group.

Method

We examined the DPA guidelines using as benchmarks two OECD Recommendations: the 2013 "Recommendation of the Council concerning Guidelines governing the Protection of Privacy and Transborder Flows of Personal Data" (OECD Privacy Guidelines), and the 2016 "Recommendation of the OECD Council on Health Data Governance" (OECD Recommendation on Health Data Governance) (1)(2). By combining the relevant principles and recommendations from these two OECD instruments, we obtained eight minimum standards against which we compared the guidance documents. These are: Data Collection and Storage; Purpose Specification and Informed Consent; Openness and Transparency; Sharing and Access to Data; Individual Participation and Data Subject Rights; Security Safeguards; Accountability. We then conducted a qualitative thematic analysis of the guidance documents to identify similarities and differences as well as possible gaps. In the area of mobile apps and privacy recent reports have noted that there is still a serious gap between legal requirements and the translation of these requirements into practical solutions. We thus further analysed only those guidelines that provided specific interpretative guidance (namely those issued in Australia, Canada, Germany, France, New Zeeland and Switzerland) and compared them against best practices listed in the UK DPA

document, which was judged as providing the most comprehensive and detailed practical advice (Appendix 2 Chapter 8).

Similarly, we screened the guidance document issued by the UK health authorities and extracted a list of criteria used to assess health apps. These are: Reference to Existing Guidelines; Effectiveness; Clinical Safety; Data Protection and Privacy; Data Security; Usability and Accessibility; Interoperability; Technical Functionality. We used this list to examine the documents developed by the EC and WHO respectively.

References

1. Organisation for Economic Co-operation and Development. The OECD Privacy Framework. 2013 [Available from: https://www.oecd.org/sti/ieconomy/oecd_privacy_framework.pdf.
2. Organisation for Economic Co-operation and Development. Recommendation of the OECD Council on Health Data Governance 2016 [Available from: https://www.oecd.org/health/health- systems/Recommendation-of-OECD-Council-on-Health- Data-Governance-Booklet.pdf.

Appendix 2 Chapter 8: Operationalising guidance across data protection authority documents

Principles	Best practice	Countries						
		Australia	Canada	Germany	France	New Zealand	Switzerland	UK
Data collection and storage limitation	Collect and process the minimum personal data necessary to the purpose	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Retain data only for the necessary time in relation to the purpose	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Purpose specification and informed consent	Provide users with information on what is being collected, used and disclosed about them and for what purpose	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Seek to obtain meaningful consent from users	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Seek consent from users to any changes that could impact on their privacy	Yes	Yes	Yes	No	Yes	No	Yes
Openness and transparency	Provide clear and informative notices	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sharing and access to data	Inform users in case data are used for marketing or advertising purposes	Yes	No	No	No	No	No	Yes
	Provide user with information on any third-party data sharing practices	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual participation and data subject rights	Pay particular attention to some groups (disabled people, children)	No	No	Yes	No	No	No	Yes
	Allow users to opt out	Yes	Yes	Yes	No	Yes	Yes	Yes
	Give users the ability to delete the data collected about them	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Security safeguads	Anonymise or de-identify personal information	Yes	No	Yes	No	Yes	Yes	Yes
	Use encryption	Yes	Yes	Yes	No	Yes	Yes	Yes
Accountability	Conduct a privacy impact assessment	Yes	No	No	No	Yes	Yes	Yes

Appendix 3 Chapter 8: Characteristics of analysed guidance documents

					Types of guidance			Format							Types of data or information mentioned		
	Publication date	**Document length (pages)**	**Specific guidance on medical and health apps**	**Stakeholders addressed**	Orientation	Interpretative	Assessment	Checklist	Questions	Good practice list	Examples	Visuals	Links to additional legal guidance	Glossary included	Personal	Sensitive	Health
Guidance from data protection authorities																	
Australia[1]	2014	21	No	App developers	Yes	Yes	No	Yes	Yes	Yes	Yes	No	Yes	No (but definitions available)	Yes	Yes	Yes
Canada[2]	2012	12	No	App developers	Yes	Yes	No	Yes	Yes	Yes	Very limited	No	Yes	No (but definitions available)	Yes	Yes	No
France[3]	2018	2	Yes	App developers	Yes	Yes	No	No	Yes	No	Very limited	No	Yes	No (but definitions available)	Yes	Yes	Yes
Germany[4]	2014	33	No	App developers	Yes	Yes	No	No	No	No	Yes	No	Yes	No	Yes	Yes	Yes
Italy[5]	2019	3	No	App users	Yes	No	No	No	Yes	No	Very limited	Video material	Yes	No	Yes	No	No
New Zealand[6]	2014	15	No	App developers	Yes	Yes	No	No	No	Yes	No	No	Very limited	No	Yes	Yes	No
Switzerland[7]	2018	15	No	App developers, public bodies, and decision makers	Yes	Yes	No	Yes	No	Yes	Yes	No	Yes	No	Yes	Yes	No
UK[8]	2013	23	No	App developers	Yes	Yes	No	No	No	Yes	Yes	Yes	Yes	No (but definitions available)	Yes	Yes	No
USA[9]	2016	12	Yes	App developers	Yes	No	No	Yes	Yes	No	No	No	Yes	Yes	No	No	Yes
Guidance from health authorities																	
Haute Autorité de Santé (France)[10]	2016	60	Yes	Manufacturers and evaluators	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
National Health Service (UK)[11]	2018	38	Yes	App developers	Yes	Yes	Yes	Yes	Yes	No	Yes	No	Yes	No (but definitions available)	Yes	Yes	Yes
Guidance from international bodies																	
Privacy Code of Conduct (European Commission)[12]	2016	23	Yes	App developers	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	No (but definitions available)	Yes	Yes	Yes
mHealth Evidence Reporting Assessment (WHO)[13]	2016	10	Yes	App developers	Yes	Yes	Yes	Yes	No	Yes	Yes	No	No	No	Yes	No	Yes

References

1. https://www.oaic.gov.au/resources/agencies-and-organisations/guides/guide-for-mobile-app-developers.pdf
2. https://www.priv.gc.ca/en/privacy-topics/technology-and-privacy/mobile-devices-and-apps/gd_app_201210/
3. https://www.cnil.fr/en/node/24293
4. https://www.bfdi.bund.de/SharedDocs/Publikationen/Entschliessungssammlung/DuesseldorferKreis/OHApps.html
5. https://www.garanteprivacy.it/temi/app
6. https://www.privacy.org.nz/assets/Files/Guidance/Need-to-Know-or-Nice-to-Have-Making-app-privacy-your-competitive-advantage.pdf
7. https://dsb.zh.ch/internet/datenschutzbeauftragter/de/publikationen/anleitungen/_jcr_content/contentPar/form_1/formitems/entwicklung_datensch/download.spooler.download.1532337698677.pdf/Leitfaden-Entwicklung-datenschutzkonformer-Apps.pdf
8. https://ico.org.uk/media/for-organisations/documents/1596/privacy-in-mobile-apps-dp-guidance.pdf
9. https://www.ftc.gov/tips-advice/business-center/guidance/mobile-health-apps-interactive-tool
10. https://www.has-sante.fr/portail/upload/docs/application/pdf/2017-03/dir1/good_practice_guidelines_on_health_apps_and_smart_devices_mobile_health_or_mhealth.pdf
11. https://developer.nhs.uk/wp-content/uploads/2018/09/Digital-Assessment-Questions-V2.1-Beta-PDF.pdf
12. http://ec.europa.eu/newsroom/dae/document.cfm?action=display&doc_id=16125
13. https://www.bmj.com/content/352/bmj.i1174

Appendix 1 Chapter 9: Characteristics of DCT systems in selected European countries.

	SWITZER-LAND	ITALY	GERMANY	IRELAND	NETHER-LANDS	UNITED KINGDOM SCOTLAND	UNITED KINGDOM ENGLAND & WALES	FRANCE
Total COVID-19 cases§	426,199	2,038,759	1,640,858	85,394	754,171	2,256,009		2,507,532
Cumulative prevalence per 1 million population§	49,245,25	33,719,77	19,584,4	17,293,99	44,013,81	33,232,31		38,415,77
Total deaths§	6,508	71,620	29,778	2,200	10,974	70,405		62,197
App name	SwissCovid	Immuni	Corona-Warn-App	COVIDTracker	CoronaMelder	ProtectScotland	NHS Covid-19	TousAntiCovid
Release date	25.05.20	15.06.20	16.06.20	07.07.20	17.08.20	14.09.20	24.09.20	22.10.20
N. of downloads	2,863,858 a	10,072,742 b	24,200,000 c	2,700,000 d	4,330,264 e	1,739,806 f	20,739,925 g	11,897,809 h
Developed in public-private partnership	✓	✓	✓	✓	✓	✓	✓	✓
Voluntariness	✓	✓	✓	✓	✓	✓	✓	✓
De-centralized protocol	✓	✓	✓	✓	✓	✓	✓	✗
Exposure parameters	1.5 meters for 15 minutes	Less than 8 meters 10 minutes	2 meters for 15 minutes	2 meters for 15 minutes	"near" for 15 minutes	2 meters for 15 minutes	2 meters for 15 minutes	2 meters for 15 minutes
Data retention:								
Random ID	14 days	14 days	14 days	14 days	14 days	14 days	14 days	14 days
Exposure code	14 days	14 days	21 days	14 days	21 days	14 days	14 days	14 days

§ at 28.12.2020 source: https://covid19.who.int

[a] at 21.12.2020 source: https://www.experimental.bfs.admin.ch/expstat/en/home/innovative-methods/swisscovid-app-monitoring.html

[b] at 28.12.2020 source: https://www.immuni.italia.it/dashboard.html

[c] at 17.12.2020 source: https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/WarnApp/Archiv_Kennzahlen/Kennzahlen_18122020.pdf?__blob=publicationFile

[d] at 28.12.2020 source: CovidTracker App

[e] at 23.12.2020 source : https://github.com/minvws/nl-covid19-notification-app-statistics/blob/main/statistics/appstore_statistics.csv

[f] at 28.12.2020 source: ProtectScotland App

[g] at 16.12.2020 source: https://www.gov.uk/government/publications/nhs-test-and-trace-england-statistics-10-december-to-16-december

[h] at 28.12.2020 source: TousAntiCovid App

Agata Ferretti

CONTACTS

Health Ethics and Policy Lab
Department of Health Sciences and Technology
Chair of Bioethics
ETH Zürich

Hottingerstrasse 10 (HOA), 8092, Zürich, Switzerland
+41 (0) 446324481
agata.ferretti@hest.ethz.ch
@FerrettiAgata

AREAS OF EXPERTISE

- Ethics and policy of health technologies
  - Artificial Intelligence
  - Medical/Health Apps
  - E-Health
  - Wearable Technologies
- Big Data Ethics and Governance
- Research Ethics & Bioethics
- Global Health Policy
- Philosophy of Science and Technology
- Neuroethics and Neuroenhancement

PROFESSIONAL EXPERIENCE

PhD. CANDIDATE & SCIENTIFIC ASSISTANT — 09/2017 – now
Health Ethics and Policy Lab, Department of Health Science and Technology, ETH Zurich, Switzerland
Supported by the Swiss National Science Foundation – NRP 75 Programme on Big Data

SCIENTIFIC ADVISER — 03/2016 – 07/2016
Sasso-Corbaro Foundation - Via Lugano 4B, 6500 Bellinzona, Switzerland

INTERN — 06/2015 – 08/2015
Global Bioethics Initiative (GBI) – 777 UN Plaza, New York, USA

EDUCATION

M.Sc. IN GLOBAL HEALTH — 2016 – 2017
London School of Economics and Political Sciences (LSE), United Kingdom
Thesis title: "*A systematic review of mobile apps effectiveness for common mental health disorders: policy implications and future recommendations*"

M.A. IN PHILOSOPHICAL SCIENCES AND BIOETHICS — 2014 – 2016
Università degli Studi di Milano, Italy
Thesis title: "*Human enhancement across science, emerging technologies and ethics*"

B.A. IN PHILOSOPHY — 2011 – 2014
Università degli Studi di Milano, Italy
Thesis title: "*Free will and responsibility. Neuroscientific discoveries challenge the freedom of the will and moral responsibility*"

TRAINING IN BIOETHICS

ERASMUS MUNDUS MASTER IN BIOETHICS — 2015 – 2016
KU Leuven, Belgium

THEORIES AND APPLICATIONS IN CONTEMPORARY ETHICS WORKSHOP — 2016
Fordham University's Rose Hill Campus, New York, USA
Recipient of the 2016 "*Fordham/Santander Universities International Scholarship in Ethics Education*"

INTERNATIONAL BIOETHICS SUMMER SCHOOL 2015
Global Bioethics Initiative, New York, USA

ZURICH WINTER SCHOOL ON THE ETHICS OF HUMAN ENHANCEMENT 2015
University of Zürich, Switzerland

HELSINKI SUMMER SCHOOL: INTRODUCTION TO BIOETHICS 2014
Helsinki University, Finland

PEER REVIEWED PUBLICATIONS

- **Ferretti, A.,** Ienca, M., Rivas Velarde, M., Hurst, S., Vayena, E. (submitted). The Challenges of Big Data for Research Ethics Committees: A Qualitative Swiss study.
- **Ferretti, A.**, Hedrich, N., Lovey, T., Vayena, E., Schlagenhauf, P. (submitted). Mobile apps for travel medicine and ethical considerations: a systematic review.
- Blasimme, A., **Ferretti, A.**, Vayena, E. (submitted). Digital contact tracing against COVID-19 in Europe: current features and ongoing developments.
- **Ferretti, A.**, Ienca, M., Sheehan, M., Blasimme, A., Dove, E. S., Farsides, B., Friesen, P., Kahn, J., Karlen, W., Kleist, P., Liao, S.M., Nebeker, C., Samuel, G., Shabani, M., Rivas Velarde, M., Vayena, E. (2021). Ethics review of big data research: what should stay and what should be reformed? *BMC Medical Ethics*. doi: https://doi.org/10.1186/s12910-021-00616-4
- **Ferretti, A.**, Ienca, M., Hurst, S., Vayena, E. (2020). Big Data, Biomedical Research, and Ethics Review: New Challenges for IRBs. *Ethics & Human Research*, doi: https://doi.org/10.1002/eahr.500065
- Jobin, A., **Ferretti, A.** (2019), The Future of Digital Health Ethics. *Bioethica Forum*, 12(1/2)
- **Ferretti, A**., Ronchi, E., Vayena, E. (2019). From principles to practice benchmarking government guidance on health apps. *The Lancet Digital Health*, doi:10.1016/S2589-7500(19)30027-5
- **Ferretti, A.**, & Ienca, M. (2018). Enhanced Cognition, Enhanced Self? On Neuroenhancement and Subjectivity. *Journal of Cognitive Enhancement*, doi: 10.1007/s41465-018-0109-9
- **Ferretti, A**., Schneider, M., Blasimme, A. (2018). Machine Learning in Medicine. *European Data Protection Law Review*, 4(3), doi:10.21552/edpl/2018/3/10
- Ienca, M., **Ferretti, A.**, Hurst, S., Puhan, M., Lovis, C., Vayena, E. (2018). Considerations for ethics review of big data health research: A scoping review. *PLOS ONE*, 13(10), e0204937, doi:10.1371/journal.pone.0204937

BOOK CHAPTERS

- Vayena, E., **Ferretti, A**. (2021). Big Data and Artificial Intelligence for Global Health. Ethical Challenges and Opportunities. In *Global Health. Ethical Challenges.* Edited by Benatar, S., Brock, G., Cambridge University Press

COMMISSIONED REPORTS

- Vayena, E., Ienca, M., Scheibner, J., **Ferretti, A**., Gille, F., Amann, J., Sleigh, J. Blasimme, A. (2019). How the General Data Protection Regulation changes the rules for scientific research. *European Parliamentary Research Service. Scientific Foresight Unit (STOA)*, EPRS_STU(2019)634447_EN.pdf

MEDIA OUTREACH & INTERVIEWS

- Interview: *The "black box" problem*. The Globe, ETH Magazine (September 2020).
  Available at: https://t.co/CYHFQpRxCP?amp=1
- Co-creator of the webpage: *Covid-19 Ethical Issues* (April 2020).
  Available at https://covid19ethics.hest.ethz.ch

CONFERENCES ATTENDED AS SPEAKER

IMPLEMENTING PERSONALISED MEDICINE RESEARCH INTO PRACTICE:
ETHICAL,LEGAL AND SOCIETAL ASPECTS (ELSA) PERSPECTIVE 11/2020
Virtual 3rd EULAC PerMed Summer School

Invited Lecture: "*Ethical Implications of Big Data and Artificial Intelligence for Personalised Medicine: an International Perspective*"

CONCEPTUAL ISSUES IMPEDING USES & SHARING OF BIG DATA IN BIOMEDICINE 11/2020
Online Webinar, Centre for Biomedical Ethics (CBmE), National University of Singapore
Invited panellist: "Which informed consent for big data research projects?"

NEW PERSPECTIVES ON THE ETHICS OF HUMAN ENHANCEMENT 02/2020
Bochum, Germany
Invited presentation: "*Is more always better? On the phenomenological implications of neuroenhancement*"

EACME CONFERENCE 2019: RETHINKING ETHICS IN 21ST CENTURY EUROPE 09/2019
Oxford, United Kingdom
Talk: "*Health-related big data research: what are the implications and recommendations for Ethical Review Committees?"*

OXFORD GLOBAL HEALTH AND BIOETHICS INTERNATIONAL CONFERENCE 2019 07/2019
Oxford, United Kingdom
Talk: "*How can AI and Big Data support the achievement of SDG#3 in resource-poor settings?*"

EUROPEAN CONFERENCE ON PHILOSOPHY OF MEDICINE AND HEALTH CARE 08/2019
Lisbon, Portugal
Talk: "*Shedding light on the black-box*"

POSTGRADUATE BIOETHICS CONFERENCE 2018 07/2018
London, United Kingdom
Talk: "*Big data trends in biomedical research: implications for ethics review boards*"

WOMEN IN BIG DATA 06/2018
Zurich, Switzerland
Poster Presentation: "*Health-related big data: challenges and implications for ethics review committees*"

INTERNATIONAL DIGITAL HEALTH CONFERENCE 04/2018
Lyon, France
Poster Presentation: "*Big data challenges in biomedical research and their implications for ethics review committees: a scoping review*"

TEACHING EXPERIENCE

- Lecturer: THE IMPACT OF DIGITAL LIFE ON SOCIETY
  ETH Zurich, Department of Health Sciences & Technology — Spring/Fall semesters 2019-2021
- Co-designer of the course: ETHICS IN MEDICINE & HEALTHCARE
  ETH Zurich, Department of Health Sciences & Technology — Spring semesters 2019/2020

SCIENTIFIC EVENTS COORDINATOR

BIG DATA CHALLENGES FOR ETHICS REVIEW COMMITTEES 04/2020
International Experts Webinar
Hosted by the Health Ethics and Policy Lab, ETH Zurich, Switzerland

TOWARDS A TRUSTWORTHY DATA GOVERNANCE FOR HEALTHCARE TECHNOLOGIES 09/2020
International Experts Webinar
Hosted by the Health Ethics and Policy Lab, ETH Zurich, Switzerland & the Centre for Biomedical Ethics, National University of Singapore

LANGUAGES

- ITALIAN – Native Speaker
- ENGLISH – High Proficiency
- FRENCH – Proficiency

JOURNAL REVIEWER

- Big Data and Society
- Digital Health
- Empirical Research on Human Research Ethics
- History and Philosophy of the Life Sciences
- Journal of Medical Ethics
- Public Health Genomics
- Trends in Neuroscience