



Article scientifique

Article

2021

Published version

Open Access

This is the published version of the publication, made available in accordance with the publisher's policy.

---

## The Challenges of Big Data for Research Ethics Committees: A Qualitative Swiss Study

---

Ferretti, Agata; Ienca, Marcello; Rivas Velarde, Minerva; Hurst, Samia; Vayena, Effy

### How to cite

FERRETTI, Agata et al. The Challenges of Big Data for Research Ethics Committees: A Qualitative Swiss Study. In: Journal of empirical research on human research ethics, 2021, p. 155626462110535. doi: 10.1177/15562646211053538

This publication URL: <https://archive-ouverte.unige.ch/unige:157122>

Publication DOI: [10.1177/15562646211053538](https://doi.org/10.1177/15562646211053538)

# The Challenges of Big Data for Research Ethics Committees: A Qualitative Swiss Study

Journal of Empirical Research on  
Human Research Ethics  
1–15

© The Author(s) 2021



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/15562646211053538

journals.sagepub.com/home/jre



Agata Ferretti<sup>1</sup> , Marcello Ienca<sup>1,2</sup>, Minerva Rivas Velarde<sup>3</sup>,  
Samia Hurst<sup>4</sup>, and Effy Vayena<sup>1</sup>

## Abstract

Big data trends in health research challenge the oversight mechanism of the Research Ethics Committees (RECs). The 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 members of the seven Swiss Cantonal RECs via semi-structured qualitative interviews. Our interviews reveal limited experience among REC members with the review of 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.

## Keywords

big data, research ethics, ethics, IRBs, biomedical research, responsible research

## 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 (Leonelli, 2020). 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 this paper, we refer to “big data research” as any research relying on large datasets, made of data heterogeneous in source, processed at high speed, and analyzed through novel computational techniques (Ienca et al., 2018).

In parallel with these changes in research practice, high profile cases of data misuse have emerged, exposing research participants to privacy breaches and risk of harm (Fuller, 2019). 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 (Ferretti et al., 2020; Rennie et al., 2020). RECs, also known as Institutional Review Boards (IRBs) and Research Ethics Boards (REBs), were created in the 20th century to protect the safety and interests of human participants in research (Friesen et al., 2019). 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 (Vayena et al., 2016).

In fact, big data research is characterized by novel ethical concerns, which can challenge traditional ethics oversight mechanisms and practices (Samuel et al., 2021).

Particularly in the biomedical and health fields, the increasing availability of digital health technologies enables the collection of an unprecedented amount of data (Car et al., 2019). 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 (Blasimme & Vayena, 2019). The potential of AI in medicine has emerged even more clearly during the COVID-19 pandemic, as differently structured data from heterogeneous sources were collected and processed for public health purposes, such as

<sup>1</sup>Health Ethics and Policy Lab, Department of Health Sciences and Technology, ETH Zürich, Switzerland

<sup>2</sup>College of Humanities, Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland.

<sup>3</sup>Department of Radiology and Medical Informatics, Faculty of Medicine, University of Geneva, Switzerland

<sup>4</sup>Institute for Ethics, History, and the Humanities, Faculty of Medicine, University of Geneva, Switzerland

## Corresponding Author:

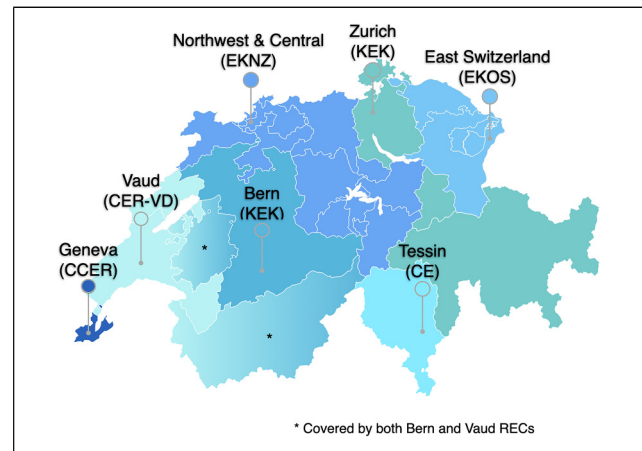
Marcello Ienca, College of Humanities, Ecole Polytechnique Fédérale de Lausanne (EPFL), CM 2 275 (Centre Midi), Station 10, CH-1015 Lausanne, Switzerland.

Email: marcello.ienca@epfl.ch

containment, mitigation, and vaccine development (Murray et al., 2020; Ngan & Kelmenson, 2020). 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, blog posts) collected through a variety of tools (e.g., social media, wearable devices) (Vayena & Gasser, 2016). 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 (Price & Cohen, 2019; Rivas Velarde et al., 2020). Second, they introduce new epistemic challenges related to the assessment of scientific validity, technological reliability, accountability, fairness, and transparency (Friesen et al., 2021). Finally, they challenge the very notion of human participants in research, as they enable retrospective data processing without physical interaction with research participants (Metcalf & Crawford, 2016).

Several questions arise about whether existing regulatory and ethical governance tools, and current practices and expertise of RECs, are adequate to protect human participants and enable ethical research (Ferretti et al., 2021). While some authors argue that the ethical principles and frameworks that traditionally govern research need to be adapted considering new research contexts (Parasidis et al., 2019; Vayena & Blasimme, 2018), studies investigating the perspectives and needs of the involved stakeholders remain scarce. Recent studies (Favaretto et al., 2020b; Samuel et al., 2021) 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 US, interviewing REC members about the ethics of social media research and research using pervasive sensing technologies (Hibbin et al. 2018; Nebeker et al., 2017; Samuel et al., 2018). 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 (Vayena, 2021).

In Switzerland, research involving human subjects, biomedical data, and biological samples requires the approval of the REC. Most of the research projects conducted in biomedical and health fields are reviewed by Cantonal RECs (Coordination Office for Human Research, 2019). Switzerland counts seven of these committees organized under Swissethics, the association of



**Figure 1.** Distribution of Swiss Research Ethics Committees (RECs) in the Swiss territory and areas of authority.

Cantonal RECs (Swissethics, 2021). 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 in a specific geographical area of Switzerland: two in the French-speaking region, one in the Italian-speaking region, and four in the German-speaking region (Figure 1). While the HRA sets general standards about RECs' composition, members' requirements, and review procedures, each REC is organized and managed independently at the Cantonal level. Although the number of members varies across committees, RECs usually include a chair, vice-chair, managing director, and scientific secretary (Swiss Federal Council, 2021).

Typically, research involving anonymous health data or biological samples is not subject to the HRA. Similarly, studies without direct implications for "the understanding of human diseases; the structure and function of the human body; or public health" (Swiss Federal Office of Public Health (FOPH), 2011) are exempted. As a consequence, human subject research in the fields of psychology, sociology, or marketing is exempted from HRA provisions. Several universities have introduced institutional ethics committees to review research projects falling outside the Cantonal RECs' purview. Nevertheless, the implementation of such intra-institutional local ethics committees is uneven throughout the country, as federal law only provides for the establishment of Cantonal RECs, and universities have no legal obligation to introduce these committees.

While a recent study looked at the experience of Swiss researchers when submitting big data research for ethical review (Favaretto et al., 2020a), no study to date has investigated the opinions and perspectives of Cantonal RECs. Therefore, this study aims to fill this gap, complement existing research, and expand knowledge on the topic by engaging with members of Cantonal RECs. Their direct experience in evaluating big data projects can provide

valuable insight into the current primary ethical oversight mechanism in Switzerland, shed light on existing gaps in the mechanism, and pave the way for needed reforms.

## Methods

### Recruitment and Sampling

For each Swiss Cantonal REC, we interviewed the chairperson (or vice-chairperson or managing director) and, whenever possible, one scientific secretary. Committees were identified through the Swissethics website. The invitation sent to each chairperson 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 email and participated in our study. Prior to recruitment, we obtained approval to conduct this study from the responsible REC.

### Interviews

Between October 2018 and May 2019, MI, AF, and MRV conducted semi-structured interviews, either face-to-face or via telephone. After written and verbal consent, each interview was recorded, and lasted between 35 min and 1 h. Interviewees could specify their preference for the interview language (French, German, Italian, English, or a combination.) We completed a total of seven interviews with 13 interviewees. Across RECs, interviewees shared similar disciplinary backgrounds (Table 1).

MI developed the interview guide (Appendix 1), which was vetted by AF and EV and approved by the research team. These interviews aim to investigate the perspective of the Cantonal RECs on (1) how to define big data research; (2) their experience with reviewing big data projects and with the ethical guidelines used for the assessment of big data research; (3) the peculiarities of big data research, namely, its benefits and challenges; and (4) the needs of RECS in order to adequately address big data research challenges (e.g., high-level recommendations, procedural good practices, education, training).

### Analysis

We transcribed *verbatim* the audio files in the original language of the interviews with the support of *Sonix* online software. Three interviews were in English, three in German, and one in Italian. To increase data consistency and reduce selective bias, we translated the non-English transcriptions into English with the assistance of *DeepL Pro* online software and additional human review. AF, MI, and MRV thematically coded and analyzed the data with *NVivo 11* Software.

Table 1. Demographics.

Number of interviews	Number of interviewees	Interviewees' Gender		Interviewees' role			Interviewees' fields of expertise *						
		Female	Male	Chair	Vice-chair	Managing director	Scientific secretary	Biology	Medicine	Pharmacology	Public health	Law/health law	Statistics
7	13	7	6	4	2	1	6	6	6	3	2	2	1

\*Each interviewee can be expert in more than one field.

Each interview was coded independently by two researchers using a combination of inductive and deductive reasoning for theme development (Fereday & Muir-Cochrane, 2006). While the deductive analysis traced the themes listed in the interview guide, the inductive analysis allowed the expansion of the list of themes, by adding those that emerged from coding the interview content.

The data analysis was performed in two steps. First, major themes of interest were identified and categorized (please refer to the Results section). This phase was duplicated by two researchers, and any disagreement was resolved with 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.

## Results

Our analysis identified four recurrent themes and several sub-themes, which are summarized in Table 2. These themes mirror the research questions addressed by this research, namely, (1) what is RECs' understanding of the "big data" concept? (2) What is the ethics review process currently in place for big data research? (3) What are RECs' perspectives about the benefits and challenges of big data research? (4) What are RECs' needs in the big data era?

### Characteristics of Big Data Research

The interviewees displayed variation in interpreting the concept of big data research, often deviating from the definition proposed in the introduction. Consistent with existing literature on the topic (Ienca et al., 2018; Jin et al., 2015), Interviewee 10 observed "there is much talking about big

data but no unanimous definition." The majority of interviewees mentioned the three versus (volume, variety, and velocity) characterizing big data, particularly stressing 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)*

Interviewees seemed aware of the diverse data sources used today for health research purposes: most mentioned data collected through social media, loyalty cards, tracking technologies, and digital health tools (e.g., health insurance mobile apps or fitness devices). Only a minority associated the big data concept with the deployment of novel analytic tools such as algorithms and 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 the re-uses of existing datasets as signals of big data research.

Although interviewees could formulate definitions of big data research, they were confused about 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 referred to projects using data and samples from biobanks. Others spoke about projects focused on improving personalized medicine, using data from tracking devices and wearables and from social media (i.e., Facebook and Twitter).

*When you talk to me about big data my idea goes more to data-bases, or biological sample banks that collect a huge amount of data and for which there is no purpose. [...] I think big data*

**Table 2.** Overview of Interview Themes and Subthemes.

Themes	Subthemes
(1) Characteristics of big data research	<ul style="list-style-type: none"> <li>Variation in big data research definitions</li> <li>Examples of big data research</li> </ul>
(2) REC mechanism in big data research	<ul style="list-style-type: none"> <li>Frequency of big data research review</li> <li>RECs' limited oversight mandate</li> <li>Criteria and guidelines used to review big data research</li> <li>Exceptionalism of big data research</li> </ul>
(3) Implications of big data research	<ul style="list-style-type: none"> <li>Benefits</li> <li>Challenges</li> </ul>
(4) RECs' needs in big data research	<ul style="list-style-type: none"> <li>Training needs</li> <li>Procedural needs</li> <li>Regulatory needs</li> </ul>

Note. REC = Research Ethics Committee.

*means analyzing a huge amount of data from various sources but without a precise purpose in mind. (Interviewee 8)*

## Current State of Ethical Oversight in Big Data Research

Six of seven Cantonal RECs reported previously reviewing and assessing big data projects. Nevertheless, our respondents emphasized that, so far, this had occurred rarely, only a few times a per year. Moreover, none of these studies were explicitly labeled by researchers as a big data project. Interviewees acknowledged their limited experience in reviewing big data research and speculated that this is because 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 from electronic health records. Furthermore, interviewees highlighted the limits of their oversight power 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." Thus this happens only rarely, because 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)*

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 (Sims, 2010), Emanuel framework (Emanuel et al. 2004), and

Beauchamp's four-principle approach (Beauchamp, 2007)), independent of the study type. RECs' assessment includes the evaluation of data protection safeguards, strategies to respect participant autonomy (i.e., informed consent), risk-benefit assessments, research purposes and data proportionality, and the scientific validity of 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)*

Our interviews revealed diverging opinions about whether the lack of specific guidance for big data research is potentially problematic. Interviewee 3 explained that the absence of such guidance 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 and project types:

*I think for that what we are seeing at the moment... 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)*

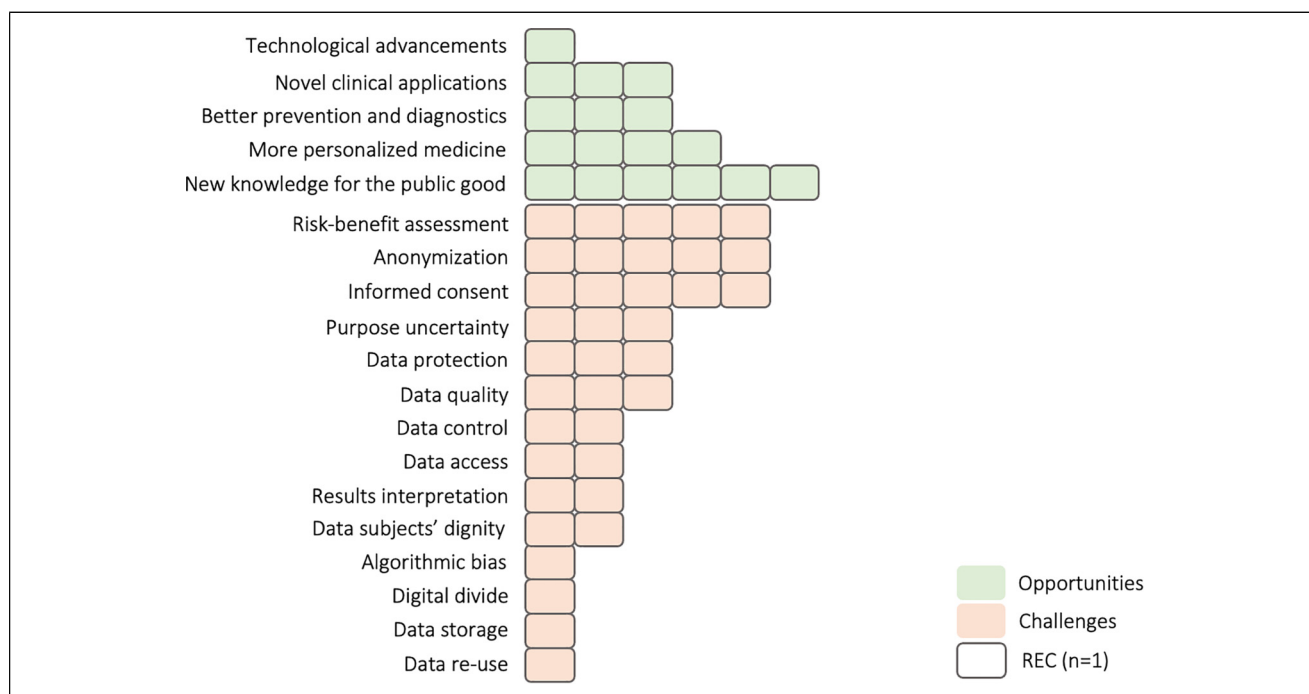
Other interviewees agreed and spoke about the HRA, Swiss data protection law, GDPR, and Emanuel framework for biomedical research as sufficient tools to guide their judgment when reviewing projects. Two interviewees openly rejected the concept of big data research exceptionalism:

*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, but diverges not only in terms of data volume and sources, but also in the types of risk involved. To demonstrate this point, Interviewee 1 commented on the potential unforeseeable risks emerging from the deployment of opaque algorithms such as deep learning to find correlations in the data:

*...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)*



**Figure 2.** Benefits and challenges of big data research discussed by Swiss Cantonal Research Ethics Committee (RECs).

In addition, Interviewee 10 spoke about the difficulty of assessing data quality in big data research compared with traditional biomedical research. If traditional research data were collected inside hospitals by researchers and health professionals, 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... 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... 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. (Interviewee 9)*

### Implications of Big Data Research

Overall, our respondents indicated a variety of benefits and challenges associated with big data research (summarized in Figure 2).

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 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. Furthermore, they 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. Interviewee 4 noted that while research participants and patients might not take direct advantage from big data research, 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)*

When asked about the challenges of big data research and their implications for ethical oversight, the respondents



identified a wide range (Figure 2). Our interviews, however, revealed a lack of consensus among Cantonal RECs concerning which challenges are most pressing (“informed consent” (Interviewee 13), “anonymization” (Interviewee 5), “results interpretation and generalizability” (Interviewee 7)). Despite this divergence, 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)). Thus, respondents stressed the importance of the rigorous application of data protection governance and the implementation of precautionary measures (e.g., data encryption and anonymization) to secure sensitive information. However, some respondents questioned the effectiveness of data protection regulations and practices 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's 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 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. 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, in light of increased privacy risks. 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:

*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 commented on the need for a dynamic form of consent. Responsible big data research should allow participants to choose for which purposes their data are used:

*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. (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), let alone for data reuses or retrospectively. Therefore, researchers should focus on obtaining consent only when collecting sensitive data. When using other data types (such as data publicly available online), researchers should rely instead on a consent waiver:

*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 of balancing 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 and the numerous possibilities for data linkage make anticipating the risks very complex (“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)). 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 control 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 AI 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 big data research makes determining fair distribution of benefits 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 biomedical big data projects *“with big data research we simply do not have the know-how” (Interviewee 6)*. 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:

*New algorithms with artificial intelligence... 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, the majority acknowledged several limitations of the oversight mechanism. When asked about their needs and envisioned solutions, REC members discussed possibilities at the levels of training, procedures, and regulations.

Regarding the first, almost all respondents expressed an urgent need to fill the REC expertise gap, recognizing expertise as a crucial factor in effectively fulfilling the oversight

mandate. Interviewees expressed interest in targeted trainings discussing characteristics, risks, and ethical implications of big data projects and AI applications in biomedicine. Interviewee 4 suggested conducting these trainings in a dynamic format, 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 consulting specialists (in the fields of big data analytics, computer science, and data management) 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” (Interviewee 4)*. However, the suggestion of Interviewee 6 to include technical experts into RECs was unpopular among 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 the procedural level, REC members clearly rejected the need for new high-level ethical guidelines *“Don’t draft them. That’s my main recommendation” (Interviewee 12)*. In fact, interviewees agreed on having adequate research ethics regulation and ethical principles already in place. Rather, several respondents stressed the need for implementable procedures to assess big data projects effectively. On this point, Interviewee 5 said that such procedures could *“help to understand how such a big data project should be structured.”* 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 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, and clarify 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. (Interviewee 3)*

Interviewee 12 further explained that the REC's attitude toward researchers is not intended to be that of watchdogs seeking 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:

*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 for researchers and for RECs), respondents listed a variety of bodies, including Swissethics, the Swiss Academy of Medical Sciences (SAMS), and the Central Ethic Committee (CEC), in collaboration with research institutions and experts in both ethics and science. If ethics review practices were introduced at an international level and made valid across countries, REC members would expect the World Health Organization (WHO) to formulate them.

Finally at the regulatory level, interviewees discussed whether the scope and mandate of RECs should be expanded to cover those big data projects currently 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. Political and health authorities, as well as RECs, should engage with society to define the boundaries of RECs' scope (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.

*Our role is to protect the individual and to decide what is in the interest of a society... 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 because “we live in a pluralistic society.”

The second point is of a pragmatic nature. As RECs do not have the capacity and expertise to review highly technical studies or studies outside the biomedical field, most respondents agreed on the idea of introducing specific oversight boards to assess the technical features of projects involving

big data and AI. These boards could complement—rather than substitute for—RECs and find their place 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 [...] 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. (Interviewee 10)*

Although this idea was endorsed by many, Interviewee 1 noted 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 in the era of big data. To succeed in this task, good will alone will not suffice. Rather, interviewees specified that Swiss regulators and policymakers should consider these gaps and further clarify the role of RECs among other ethical oversight mechanisms in place.

## Strengths and Limitations

While the methodology of qualitative interview analysis allows for the detailed exploration of opinions and perspectives, 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 cultural variation that exists within the country. Furthermore, since the Swiss ethical oversight mechanism partially resembles those of other European countries (e.g., neighboring Germany, Austria, France, and Italy) and internationally, a certain degree of generalization of results could be justified.

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 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, in practice, set the agenda for the committee, and scientific secretaries first review and evaluate 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 COVID-19 pandemic can be interpreted as a limitation. Indeed, the pandemic has increased pressure on RECs, especially for reviewing public health projects that leverage the power of big data and AI. Nonetheless, the results of this study transcend the temporality of current research conditions, as they relate to the complex oversight system of Cantonal RECs, which is not evolving as rapidly. Future research could explore which processes and functions of Cantonal RECs have changed as a result of the COVID-19 pandemic.

## 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 conceptual tools traditionally used to assess biomedical research appear 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. As previous studies indicated, REC members' uncertainty could result in inconsistencies across committees (Favaretto et al, 2020b; Vitak et al., 2017). Moreover, the way in which interviewees define big data can influence their assessment of the most pressing ethical challenges. By using narrow definitions of big data—namely focusing on one or few characteristics (e.g., data source, data volume)—RECs may be more sensitive to some ethical implications than others. It is relevant to note, 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 (Ienca et al., 2018). While some interviewees believed that the lack of specific ethical guidance does not negatively impact ethical review practices, others expressed concern about having to interpret and

judge big data research on a case-by-case basis without guidance. These diverging opinions might mirror different RECs' approaches to ethics review, as well as RECs' members' confidence levels, experience in managing big data projects, and expertise in the technical disciplines (i.e., data and computer sciences). RECs' diverging interpretations may result in disharmonious evaluations and decisions across committees, which could negatively affect researchers' trust in the oversight system, data sharing practices, and research collaborations (Ballantyne & Schaefer, 2020; Dove & Garattini, 2018; van den Hoonaard & Hamilton, 2016). Although a lack of transparency about evaluation procedures and inconsistencies across RECs' judgments are not exclusive to big data research (Lynch et al., 2020), our findings show 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 fundamental technical aspects characterizing these studies. 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 (von Elm & Briel, 2019). However, unless the law is amended to expand the purview mission of the ethical oversight mechanism, Cantonal RECs have no choice but to invite researchers to submit their research voluntarily. Some studies have suggested that RECs should engage researchers in a dialogue to make them aware of and accountable for the consequences of their research (Holland, 2018). Concerning RECs' insufficient expertise around technical features of big data research, our interviewees were aware of their shortcomings. Most REC members expressed both willingness and commitment to implementing strategies to overcome their weaknesses (e.g., involving data specialists in the assessment of big data research, or attending trainings to increase their technical competence). Furthermore, our findings partially align with the results of a recent study focusing on IRB staff's perspectives in the United States (Vitak et al., 2017). The IRBs surveyed believed that over time they would surmount their shortcomings in assessing the technical aspects of big data research proposals, through cumulative experience. Yet, the rapidity with which AI technology and big data applications evolve further complicates RECs' and IRBs' attempts to get up to speed in their subject matter knowledge (Nebeker et al., 2019; Prosperi & Bian, 2019).

### *Scope of Ethical Reflection*

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 which tools are best suited to address them. These different opinions among interviewees might be explained by their background, personal bias, and the lack of training in big data research. 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 regulations, may signal a problem. Some authors flag the risk of viewing these tools as ethics panacea (Babb, 2020; Corrigan, 2003). 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, for example, justice, dignity, and fairness (Ballantyne, 2019; McKeown et al., 2021). Our results highlight this gap in 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 assessing unforeseeable and novel risks (Sheehan et al., 2019). This concern, which remains unresolved for the time being, underscores the need for a broader conversation in society about the importance of big data research, and its uses in terms of our collective interest (McMahon et al., 2020).

### Proposals for Reform

Our results shed light on the limitations of the current mechanism of Cantonal RECs, in terms of skills, practices, and guidelines. REC members—aware of these shortcomings—suggested possible solutions to tackle them. The interviewees' request for training on big data and AI reveals interest in expanding their knowledge. In addition, the practice of involving experts to fill RECs' expertise gap can be seen as an attempt to offer assistance to researchers (Huh-Yoo et al., 2021). Interviewees' desire to improve the *status quo* of ethical review is further evidenced by their suggestion of creating complementary oversight mechanisms (e.g., big data boards), to 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 in private and public sectors, regardless of data types and sources. 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 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.

### Best Practice

Swiss Cantonal RECs should be reformed if they are to be effective 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 and the regulations that define the mandate of RECs. Ethics oversight mechanisms outside the Swiss context might benefit from similar revisions. In addition, this study suggests that researchers be proactive in reaching out for RECs' opinions and 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, universities, and funding bodies) (Samuel et al., 2019).

### 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 align with the literature at the international level. 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 (ÓhÉigeartaigh et al., 2020). The recent COVID-19 pandemic exemplified how divergent laws governing research, unclear ethical evaluation methods, and unrobust oversight mechanisms can slow down research processes, jeopardize efforts for public health, and reduce public trust in scientific institutions (Gardner et al., 2020).

### Educational Implications

Our results emphasize the need for knowledge exchange and a more productive engagement among the various factors involved in big data research. These include and are not limited to RECs, researchers, research institutions, private enterprises, citizen science groups, and the public (Vayena & Gasser, 2016). In particular, if on the one hand REC members should acquire more technical skills about, for example, data analysis methodologies and AI-enabled technologies, researchers should also 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 con-

frontation on diverging perspectives (Zook et al., 2017). 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.

## Appendix I: Interview guide

---

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 benefits 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 govts, private corporations, etc.)? At which level (international vs. national)?
- 

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.

Moreover, the study investigator will ask to get permission for tape recording.

## Acknowledgments

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 Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung (Grant No. 407540\_167223).


## Ethics Statement

Prior to recruitment, we obtained approval (EK 2017-N-74) to conduct this study from ETH Zurich's Research Ethics Committee.

## Authors' Contributions

MI, AF, and MRV conducted and analyzed the interviews and compiled the results. AF drafted the manuscript. All authors contributed to the study design and the development of the final manuscript and approved the submitted version.

## ORCID iD

Agata Ferretti  <https://orcid.org/0000-0001-6716-5713>

## Supplemental Material

Supplemental material for this article is available online.

## References

- Babb, S. (2020). Regulating human research. In *Regulating human research*. Stanford University Press.
- Ballantyne, A. (2019). Adjusting the focus: A public health ethics approach to data research. *Bioethics*, 33(3), 357–366. <https://doi.org/10.1111/bioe.12551>
- Ballantyne, A., & Schaefer, G. O. (2020). Public interest in health data research: Laying out the conceptual groundwork. *Journal of Medical Ethics*, 46(9), 610–616. <http://dx.doi.org/10.1136/medethics-2020-106152>
- Beauchamp, T. L. (2007). The 'four principles' approach to health care ethics. *Principles of Health Care Ethics*, 29, 3–10. <https://doi.org/10.1002/9780470510544.ch1>
- Blasimme, A., & Vayena, E. (2019). The ethics of AI in biomedical research, patient care and public health. In M. D. Dubber, F. Pasquale, & S. Das (Eds.) *Oxford handbook of ethics of artificial intelligence* (p. 718). Oxford University Press.
- Car, J., Sheikh, A., Wicks, P., & Williams, M. S. (2019). Beyond the hype of big data and artificial intelligence: Building foundations for knowledge and wisdom. *BMC Med* 17, 143. <https://doi.org/10.1186/s12916-019-1382-x>
- Coordination Office for Human Research (2019). *The Human Research Act and the Ethics Committees for Research*. Retrieved March 20, 2021, <https://www.bag.admin.ch/bag/en/home/medizin-und-forschung/forschung-am-menschen/koordinationsstelle-forschung-mensch.html>.
- Corrigan, O. (2003). Empty ethics: The problem with informed consent. *Sociology of Health & Illness*, 25(7), 768–792. <https://doi.org/10.1046/j.1467-9566.2003.00369.x>
- Dove, E. S., & Garattini, C. (2018). Expert perspectives on ethics review of international data-intensive research: Working towards mutual recognition. *Research Ethics*, 14(1), 1–25. <https://doi.org/10.1177/1747016117711972>
- Emanuel, E. J., Wendler, D., Killen, J., & Grady, C. (2004). What makes clinical research in developing countries ethical? The benchmarks of ethical research. *The Journal of Infectious Diseases*, 189(5), 930–937. <https://doi.org/10.1086/381709>
- Favaretto, M., De Clercq, E., Briel, M., & Elger, B. S. (2020a). 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*, 15(4), 339–354. <https://doi.org/10.1177/1556264620935223>
- Favaretto, M., De Clercq, E., Schneble, C. O., & Elger, B. S. (2020b). What is your definition of big data? Researchers' understanding of the phenomenon of the decade. *PLoS One*, 15(2), e0228987. <https://doi.org/10.1371/journal.pone.0228987>
- Fereday, J., & Muir-Cochrane, E. (2006). Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. *International Journal of Qualitative Methods*, 5(1), 80–92. <https://doi.org/10.1177/160940690600500107>
- Ferretti, A., Ienca, M., Hurst, S., & Vayena, E. (2020). Big data, biomedical research, and ethics review: New challenges for IRBs. *Ethics & Human Research*, 42(5), 17–28. <https://doi.org/10.1002/eahr.500065>
- Ferretti, A., Ienca, M., Sheehan, M., Blasimme, A., Dove, E. S., Farsides, B., & Kleist, P., Matthew Liao, S., 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*, 22(1), 1–13. <https://doi.org/10.1186/s12910-021-00616-4>
- Friesen, P., Douglas-Jones, R., Marks, M., Pierce, R., Fletcher, K., Mishra, A., & Graham, M. (2021). Governing AI-driven health research: Are IRBs up to the task? *Ethics & Human Research*, 43(2), 35–42. <https://doi.org/10.1002/eahr.500085>
- Friesen, P., Redman, B., & Caplan, A. (2019). Of straws, camels, research regulation, and IRBs. *Therapeutic Innovation & Regulatory Science*, 53(4), 526–534. <https://doi.org/10.1177/2168479018783740>
- Fuller, M. (2019). Big data and the facebook scandal: Issues and responses. *Theology*, 122(1), 14–21. <https://doi.org/10.1007/s11948-021-00282-0>
- Gardner, L., Ratcliff, J., Dong, E., & Katz, A. (2020). A need for open public data standards and sharing in light of COVID-19. *The Lancet Infectious Diseases*, 21(4), e80. [https://doi.org/10.1016/S1473-3099\(20\)30635-6](https://doi.org/10.1016/S1473-3099(20)30635-6)
- Hibbin, R. A., Samuel, G., & Derrick, G. E. (2018). 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*, 13(2), 149–159. <https://doi.org/10.1177/1556264617751510>
- Holland, K. (2018). 18. Enriching ethics review processes in the spirit of participatory dialogue. In A. Hamilton & W. C. van den Hoonaard (Eds.), *The ethics rupture* (pp. 353–375). University of Toronto Press.
- Huh-Yoo, J., Kadri, R., & Buis, L. R. (2021). Pervasive healthcare IRBs and ethics reviews in research: Going beyond the paper-work. *IEEE Pervasive Computing*, 20(1), 40–44. <https://doi.org/10.1109/MPRV.2020.3044099>
- Ienca, M., Ferretti, A., Hurst, S., Puhon, M., Lovis, C., & Vayena, E. (2018). Considerations for ethics review of big data health research: A scoping review. *PLoS One*, 13(10), e0204937. <https://doi.org/10.1371/journal.pone.0204937>
- Jin, X., Wah, B. W., Cheng, X., & Wang, Y. (2015). Significance and challenges of big data research. *Big Data Research*, 2(2), 59–64. <https://doi.org/10.1016/j.bdr.2015.01.006>
- Leonelli, S. (2020). *Scientific research and big data*. Retrieved May 29, 2020, from <https://plato.stanford.edu/entries/science-big-data/#BigDataRiskEthiDataScie>
- Lynch, H. F., Abdirisak, M., Bogia, M., & Clapp, J. (2020). Evaluating the quality of research ethics review and oversight: A systematic analysis of quality assessment instruments. *AJOB Empirical Bioethics*, 11(4), 208–222. <https://doi.org/10.1080/23294515.2020.1798563>
- McKeown, A., Mourby, M., Harrison, P., Walker, S., Sheehan, M., & Singh, I. (2021). Ethical issues in consent for the reuse of data in health data platforms. *Science and Engineering Ethics*, 27(1), 1–21.
- McMahon, A., Buyx, A., & Prainsack, B. (2020). 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*, 28(1), 155–182. <https://doi.org/10.1093/medlaw/fwz016>
- Metcalf, J., & Crawford, K. (2016). Where are human subjects in big data research? The emerging ethics divide. *Big Data & Society*, 3(1), <https://doi.org/10.1177/2053951716650211>
- Murray, C. J., Alamro, N. M. S., Hwang, H., & Lee, U. (2020). Digital public health and COVID-19. *The Lancet Public Health*, 5(9), e469–e470. [https://doi.org/10.1016/S2468-2667\(20\)30187-0](https://doi.org/10.1016/S2468-2667(20)30187-0)
- Nebeker, C., Harlow, J., Espinoza Giacinto, R., Orozco-Linares, R., Bloss, C. S., & Weibel, N. (2017). Ethical and regulatory challenges of research using pervasive sensing and other emerging technologies: IRB perspectives. *AJOB Empirical Bioethics*, 8(4), 266–276. <https://doi.org/10.1080/23294515.2017.1403980>
- Nebeker, C., Torous, J., & Ellis, R. J. B. (2019). Building the case for actionable ethics in digital health research supported by artificial intelligence. *BMC Medicine*, 17(1), 1–7. <https://doi.org/10.1186/s12916-019-1377-7>
- Ngan, O. M., & Kelmenson, A. M. (2020). Using big data tools to analyze digital footprint in the COVID-19 pandemic: Some public health ethics considerations. *Asia Pacific Journal of Public Health*, 33(1), 129–130. <https://doi.org/10.1177/1010539520984360>
- ÓhÉigeartaigh, S. S., Whittlestone, J., Liu, Y., Zeng, Y., & Liu, Z. (2020). Overcoming barriers to cross-cultural cooperation in AI ethics and governance. *Philosophy & Technology*, 33(4), 571–593. <https://doi.org/10.1007/s13347-020-00402-x>
- Parasidis, E., Pike, E., & McGraw, D. (2019). A Belmont report for health data. *The New England Journal of Medicine*, 380(16), 1493–1495. <https://doi.org/10.1056/NEJMp1816373>
- Price, W. N., & Cohen, I. G. (2019). Privacy in the age of medical big data. *Nature Medicine*, 25(1), 37–43. <https://doi.org/10.1038/s41591-018-0272-7>
- Prosperi, M., & Bian, J. (2019). Is it time to rethink institutional review boards for the era of big data? *Nature Machine Intelligence*, 1(6), 260–260. <https://doi.org/10.1038/s42256-019-0059-7>
- Rennie, S., Buchbinder, M., Juengst, E., Brinkley-Rubinstein, L., Blue, C., & Rosen, D. L. (2020). Scraping the web for public health gains: Ethical considerations from a ‘big data’ research project on HIV and incarceration. *Public Health Ethics*, 13(1), 111–121. <https://doi.org/10.1093/phe/phaa006>
- Samuel, G., Chubb, J., & Derrick, G. (2021). Boundaries between research ethics and ethical research use in artificial intelligence health research. *Journal of Empirical Research on Human Research Ethics*, 16(3), 325–337. <https://doi.org/10.1177/15562646211002744>
- Rivas Velarde, M. C. R., Tsantoulis, P., Burton-Jeangros, C., Aceti, M. A., Chappuis, P., & Hurst, S. (2020). Citizens’ views on sharing their health data: The role of competence, reliability and pursuing the common good. *BMC Med Ethics* 22, 62. <https://doi.org/10.1186/s12910-021-00633-3>
- Samuel, G., Derrick, G. E., & van Leeuwen, T. (2019). The ethics ecosystem: Personal ethics, network governance and regulating actors governing the use of social media research data. *Minerva*, 57(3), 317–343. <https://doi.org/10.1177/15562646211002744>
- Samuel, G., Derrick, G., & van Leeuwen, T. N. (2018). Ethical challenges around the use of social media data: Views of researchers and research ethics committee members. STI 2018 Conference Proceedings, 221–227. Centre for Science and Technology Studies (CWTS). Available at: <https://hdl.handle.net/1887/65262>
- Sheehan, M., Thompson, R., Fistein, J., Davies, J., Dunn, M., Parker, M., & Woods, K. (2019). Authority and the future of consent in population-level biomedical research. *Public Health Ethics*, 12(3), 225–236. <https://doi.org/10.1093/phe/phz015>
- Sims, J. M. (2010). A brief review of the Belmont report. *Dimensions of Critical Care Nursing*, 29(4), 173–174. <https://doi.org/10.1097/DCC.0b013e3181de9ec5>
- Swiss Federal Council (2021). *Ordinance on organisational aspects of the Human Research Act*. Retrieved March 30, 2021, from [https://www.fedlex.admin.ch/eli/cc/2013/644/en#art\\_3](https://www.fedlex.admin.ch/eli/cc/2013/644/en#art_3)
- Swiss Federal Office of Public Health (FOPH) (2011). *Swiss Federal Human Research Act (HRA)*. Retrieved from <https://fedlex.data.admin.ch/filestore/fedlex.data.admin.ch/eli/cc/2013/617/20140101/en/pdf-a/fedlex-data-admin-ch-eli-cc-2013-617-20140101-en-pdf-a.pdf>
- Swissethics (2021). *Mission statement of Swissethics*. Retrieved March 30, 2021, from <https://swissethics.ch/en/>
- van den Hoonaard, W. C., & Hamilton, A. (Eds.). (2016). *The ethics rupture: Exploring alternatives to formal research-ethics review*. University of Toronto Press.



- Vayena, E. (2021). Value from health data: European opportunity to catalyse progress in digital health. *The Lancet*, 397(10275), 652–653. [https://doi.org/10.1016/S0140-6736\(21\)00203-8](https://doi.org/10.1016/S0140-6736(21)00203-8)
- Vayena, E., & Blasimme, A. (2018). Health research with big data: Time for systemic oversight. *The Journal of Law, Medicine & Ethics*, 46(1), 119–129. <https://doi.org/10.1177/1073110518766026>
- Vayena, E., & Gasser, U. (2016). Strictly biomedical? Sketching the ethics of the big data ecosystem in biomedicine. In B. Mittelstadt & L. Floridi (Eds.), *The ethics of biomedical big data* (pp. 17–39). Law, governance and technology series, vol 29. Cham: Springer. [https://doi.org/10.1007/978-3-319-33525-4\\_2](https://doi.org/10.1007/978-3-319-33525-4_2).
- Vayena, E., Gasser, U., Wood, A. B., O'Brien, D., & Altman, M. (2016). Elements of a new ethical framework for big data research. *Washington and Lee Law Review Online*, 72(3), 420–441. <http://nrs.harvard.edu/urn-3:HUL.InstRepos:28552577>
- Vitak, J., Proferes, N., Shilton, K., & Ashktorab, Z. (2017). Ethics regulation in social computing research: Examining the role of institutional review boards. *Journal of Empirical Research on Human Research Ethics*, 12(5), 372–382. <https://doi.org/10.1177/1556264617725200>
- von Elm, E., & Briel, M. (2019). *Survey on researchers' opinion about and experience with the Swiss Federal Act on Research involving human beings*. Retrieved October 27, 2021, from [https://www.bag.admin.ch/dam/bag/de/dokumente/biomed/forschung-am-menschen/forschung-biomedizin/Hauptbericht-Befragung.pdf.download.pdf/190715\\_EDFI\\_Main%20Report.pdf](https://www.bag.admin.ch/dam/bag/de/dokumente/biomed/forschung-am-menschen/forschung-biomedizin/Hauptbericht-Befragung.pdf.download.pdf/190715_EDFI_Main%20Report.pdf)
- Zook, M., Barocas, S., Boyd, D., Crawford, K., Keller, E., Gangadharan, S. P., & Metcalf, J. (2017). *Ten simple rules for responsible big data research*. Public Library of Science San Francisco. <https://doi.org/10.1371/journal.pcbi.1005399>

## Authors Biography

**Agata Ferretti** is a Postdoctoral Researcher at the Health Ethics & Policy Lab, Department of Health Sciences and Technology at ETH Zurich, Switzerland. Her PhD research focused on the ethics and governance of big data in health research and digital health applications.

**Marcello Ienca** is a Principal Investigator at the College of Humanities at EPFL where he leads the ERA-NET funded Intelligent Systems Ethics research unit. His research focuses on the ethical, legal, and social implications of neurotechnology and artificial intelligence, with particular focus on big data trends in neuroscience and biomedicine, human-machine interaction, social robotics, digital health, and cognitive assistance for people with intellectual disabilities.

**Minerva Rivas Velarde** is an SNSF Ambizione Group Leader at the Department of Radiology and Medical Informatics, University of Geneva, Switzerland. Her research focuses on global health, eHealth, disability studies, and bioethics.

**Samia Hurst** is Professor of Bioethics, Director of the Institute of Ethics, History, Humanities (IEH2) and of the Department of Health and Community Medicine at the Faculty of Medicine of Geneva, Switzerland. She has been working on ethical issues in clinical practice, health policy ethics, particularly issues of equity and protection of the vulnerable, and ethical issues in personalized medicine.

**Effy Vayena** is Professor of Bioethics and Director of the Health Ethics & Policy Lab, Department of Health Sciences and Technology at ETH Zurich, Switzerland. Her work focuses on the important societal issues of data and technology, as they relate to scientific progress and how it is or should be applied to public and personal health.