What are the bottlenecks to health data sharing in Switzerland? An interview study

Kelly E. Ormond[a], Sabine Bavamian[b], Claudia Becherer[cd], Christine Currat[b], Francisca Joerger[ce], Thomas R. Geiger[g], Elke Hiendlmeyer[cf], Julia Maurer[h], Timo Staub[i], Effy Vayena[a]

a   D-HEST,Health Ethics and Policy Lab, ETH-Zurich, Zurich, Switzerland
b   Swiss Biobanking Platform, Epalinges, Switzerland
c   Swiss Clinical Trial Organisation, Bern, Switzerland
d   Department Clinical Research (DKF), University Basel, University Hospital Basel, Basel, Switzerland
e   Clinical Trials Center, University Hospital Zurich, Zurich, Switzerland
f   Clinical trials unit (CTU), Kantonsspital St. Gallen, St. Gallen, Switzerland
g   Swiss Personalized Health Network (SPHN), Swiss Academy of Medical Sciences, Bern, Switzerland
h   Personalised Health Informatics Group, SIB Swiss Institute of Bioinformatics, Lausanne, Switzerland
i   Bern Center for Precision Medicine, University of Bern, Bern, Switzerland

Summary

BACKGROUND: While health data sharing for research purposes is strongly supported in principle, it can be challenging to implement in practice. Little is known about the actual bottlenecks to health data sharing in Switzerland.

AIMS OF THE STUDY: This study aimed to assess the obstacles to Swiss health data sharing, including legal, ethical and logistical bottlenecks.

METHODS: We identified 37 key stakeholders in data sharing via the Swiss Personalised Health Network ecosystem, defined as being an expert on sharing sensitive health data for research purposes at a Swiss university hospital (or a Swiss disease cohort) or being a stakeholder in data sharing at a public or private institution that uses such data. We conducted semi-structured interviews, which were transcribed, translated when necessary, and de-identified. The entire research team discussed the transcripts and notes taken during each interview before an inductive coding process occurred.

RESULTS: Eleven semi-structured interviews were conducted (primarily in English) with 17 individuals representing lawyers, data protection officers, ethics committee members, scientists, project managers, bioinformaticians, clinical trials unit members, and biobank stakeholders. Most respondents felt that it was not the actual data transfer that was the bottleneck but rather the processes and systems around it, which were considered time-intensive and confusing. The templates developed by the Swiss Personalised Health Network and the Swiss General Consent process were generally felt to have streamlined processes significantly. However, these logistics and data quality issues remain practical bottlenecks in Swiss health data sharing. Areas of legal uncertainty include privacy laws when sharing data internationally, questions of "who owns the data", inconsistencies created because the Swiss general consent is perceived as being implemented differently across different institutions, and definitions and operationalisation of anonymisation and pseudo-anonymisation. Many participants desired to create a "culture of data sharing" and to recognise that data sharing is a process with many steps, not an event, that requires sustainability efforts and personnel. Some participants also stressed a desire to move away from data sharing and the current privacy focus towards processes that facilitate data access.

CONCLUSIONS: Facilitating a data access culture in Switzerland may require legal clarifications, further education about the process and resources to support data sharing, and further investment in sustainable infrastructureby funders and institutions.

Introduction

Sharing health-related data for research purposes has many potential benefits: transparency of published clinical research, the potential to validate research results, adding scientific knowledge without requiring new data collection, and honouring the public good aspects of medical research [1]. Numerous laws, ordinances and policies currently exist to regulate the sharing of confidential health data for research purposes (table 1) in Switzerland and internationally. Despite these regulations, there is room for different interpretations, and Swiss stakeholders consistently express concerns about how to share health data in a legally and ethically appropriate manner across institutions within Switzerland and across international borders.

The Swiss Personalised Health Network (SPHN) initiative by the Swiss Federal Government aims to establish coordinated infrastructures to make health data available, interoperable, and shareable for research in Switzerland. This network provides a secure platform (BioMedIT) for processing sensitive data with dedicated computational resources, maintaining research subject privacy to the greatest possible extent [2]. The Swiss Personalised Health Network Interoperability Framework turns (routine) health data into FAIR (findable, accessible, interoperable, and

Kelly E. Ormond
Health Ethics and Policy Lab
Hottingerstrasse 10
CH-8092 Zürich
kelly.ormond[at]hest.ethz.ch

reusable) research data. The Swiss Personalised Health Network facilitates a consolidated contractual framework to tackle challenges related to ethical, legal, and social issues (ELSI), providing legal agreement templates, the SPHN Ethical Framework, and the Public-Private Partnerships Guidelines [3]. Within this initiative, projects are funded with the main focus on contributing to developing a nationwide data-sharing infrastructure. However, researchers and coordinators are still confronted with ongoing technical- and ELSI-related hurdles. In the recent joint SPHN–PHRT (Personalised Health and Related Technologies, an ETH Domain program [ETH: Eidgenössische Technische Hochschule]) project call, the National Data Streams, participating parties struggled with missing legal clarifications and uncertainties of governance-related aspects and their associated intellectual property regulations to share health data with third parties.

Bottlenecks to health data sharing can significantly burden planned research projects through significant time delays or even prevent their conduct because of the existing uncertainties (for example, due to different interpretations of the law or the law itself). These bottlenecks can occur due to the views of potential participants or expert stakeholders, who might be researchers, lawyers, or institutions. While recently published studies have described the general public's views on data sharing in Switzerland [4–6], only limited data (collected in 2018) exists from the expert stakeholders [7, 8]. This study aimed to better understand the potential bottlenecks to health data sharing in Switzerland by interviewing relevant expert stakeholders at the various Swiss university hospitals and other organisations where health details are shared and used for research.

Materials and methods

Due to a lack of data on bottlenecks to data sharing explicitly in Switzerland, we chose a qualitative methodology, specifically semi-structured interviews. Interviews are an excellent method to obtain rich data, particularly when little is known about a topic, and this approach allows an in-depth exploration to better understand the nuances and reasons behind specific behaviours [9]. The study instrument and methods were approved by the ethics committee at ETH Zurich (2021-N-131).

Instrument

Because this was an exploratory study, the entire research team developed a novel semi-structured interview guide (available for download at https://doi.org/10.57187/s.3538), drawing upon their experiences in health data sharing from their perspectives as researchers, in biobanking and clinical trials, and in bioethics. The questions cov-

ered several areas, including the participant's background, general data-sharing experiences, and perception of areas that were unclear or roadblocks to data sharing. The interview guide also had specific prompts on logistics, law/regulation/policy, privacy, consent, and data reuse. Questions focused on health data sharing within Switzerland and with collaborating research teams in other countries. Importantly, we did not restrict the definition of health data but allowed our study participants to define it broadly.

Recruitment

The inclusion criteria for this study were defined as being an expert on sharing sensitive health data for research purposes at a Swiss university hospital (or a Swiss disease cohort) or being a stakeholder in data sharing at a public or private institution that uses such data, focusing on researchers who were part of the Swiss Personalised Health Network research ecosystem. Example roles included university hospital lawyers, members/chairs of hospital data access committees, hospital or university IT group members (e.g. BioMedIT nodes), clinical trial units, and principal investigators (PIs)/researchers of SPHN-funded projects. The research team identified 37 potential interviewees across Switzerland. They included data protection officers, security officers/node managers from BioMedIT nodes, and members of biobanks/registries, hospital IT, clinical trial teams, sample providers and cohorts, ethics committees, data governance boards, legal departments, and the BBMRI, the European research infrastructure for biobanking (https://www.bbmri-eric.eu/). The potential participants were sent an email invitation and asked to share it with relevant colleagues at their institutions, if appropriate, both in the recruitment stage and after any conducted interviews (snowball sampling). Reminder emails were sent six weeks after the first email. Recruitment was stopped when data saturation was reached, defined as a lack of new themes after several interviews [10].

Method

We used a semi-structured interview approach. Verbal informed consent was obtained before starting the interviews. The interviews were conducted on Zoom individually or in groups of up to three individuals and were recorded. A single interviewer conducted all but two interviews in English; the two remaining interviews were conducted in German by a native German speaker. Field notes were also taken during each interview. The audio files were transcribed using a combination of Trint, a cloud-based transcription tool that uses artificial intelligence to automatically transcribe audio recordings in multiple languages, and DeepL Translator, a neural machine translation service, to translate the German interviews into Eng-

Table 1:
Applicable regulations and guidance.

Swiss Human Research Act (SR 810.30) and its ordinances and regulations
Swiss Federal Act on Data Protection (SR 235.1) and relevant cantonal regulations
General Consent (https://swissethics.ch/templates/studieninformationen-und-einwilligungen, https://swissethics.ch/news/2019/02/22/veroeffentlichung-der-version-2-zum-nationalen-generalkonsent)
Information Security (Swiss Personalised Health Network, 2018) and Responsible Data Sharing (SPHN, 2018)
General Data Protection Regulation (GDPR) (Regulation EU, 2016)
Health Insurance Privacy Accountability Act (United States)
2018 Data Protection Act (UK)

lish. The original interviewer then checked the transcriptions for fidelity.

Analysis

Interview transcripts were manually reviewed to remove identifying information (names, institutions, and cities) and to "fact-check" the transcriptions. A single investigator (KEO) reviewed the field notes and transcripts to identify preliminary themes, which were then discussed by the entire research team. An inductive codebook was developed, and a single interviewer (KEO) used NVivo 1.7.1 (a qualitative analysis software package) to code the interview transcripts and extract the themes presented here. Quotes were selected for representativeness and edited slightly for conciseness and clarity (e.g. excluding filler words). The entire research team reviewed the edited quotes to ensure these minor edits did not alter their context and meaning.

Results

Eleven interviews, each lasting 30–65 minutes, were conducted with 17 (of 37 potential) individuals between February and April 2022. These interviewees represented five of the major universities and hospitals (Eidgenössische Technische Hochschule ETH, University Hospital Zurich, Inselspital Berne, University Hospital Basel, and Centre hospitalier universitaire vaudois (CHUV). The interviewees had a mix of experience in different roles around data-sharing; they included lawyers (n = 3), members of ethics committees or applied ethicists (n = 2), basic scientists actively involved in sharing or receiving data for their research (n = 2), data protection officers (n = 4), individuals who deal with data logistics (e.g. bioinformatics, IT, or clinical trial services; n = 4), an individual involved in project management at a large registry (n = 1), and an individual involved in biobanking oversight internationally (n = 1). The major themes are summarised in table 2, with representative quotes to illustrate the theme.

Logistics are both facilitators and barriers

Interviewees in this study described that for data sharing to occur, researchers (data providers and recipients) needed to be transparent about what kind of data was being shared, how sensitive and identifiable the data was, what the purpose of sharing was, and to describe the legal basis for the sharing (whether consent or another basis). Participants appreciated the various legal and other templates created for use within Switzerland, such as those developed by Swissethics, the Swiss Biobanking Platform (SBP), or the Swiss Personalised Health Network, and the nationwide general consent process. These were seen as streamlining and standardising logistics. While interviewees could list areas that were challenging in the data-sharing process, most felt that the actual transfer component of data sharing, or even dealing with participant withdrawals, was not a major bottleneck (table 2). Rather, the overall processes and systems around data sharing, their time intensiveness, and related confusion serve as a barrier. As such, several interviewees recommended involving data-sharing service teams early

(at the planning stage) to facilitate these logistical challenges.

General consent serves as a facilitator to Swiss data-sharing

Overall, interviewees expressed that Switzerland's approach of having general consent is seen as a facilitator to data sharing within the country (table 2). Nevertheless, some expressed worry that general consent lessens the connection between the researcher or data-providing institution and the patient (data subject). Specifically, due to its broad approach, it loses the meaning that one would otherwise have in an individual study consent, lessening trust. Two areas of general consent remained confusing to interviewees. The first was the perception that hospitals have adopted slightly different versions of general consent across the country, and this lack of harmonisation is felt to create challenges. Second, some interviewees remain unclear how often and in what circumstances research participants should be recontacted and reconsented. Finally, regarding public-private partnerships, interviewees commented that transparency is generally important since the public may have different levels of comfort with this type of data sharing. For example, research teams spinning off a company based on academic research should be especially clear that they have the correct consent for data sharing before taking the data to an industry setting since this might be considered the commercial use of health data.

Data quality issues are a bottleneck

Health data quality issues, including data interoperability, occur at several levels and appear to be a real bottleneck to sustainable data sharing. Interviewees described that it is hard to obtain high-quality data from clinical records. Neither Switzerland nor most other EU countries collect structured longitudinal health data that can be used for research purposes. Additionally, the electronic health record (EHR) is structured for clinical interactions and not research; the abundance of free text and lack of consistently adopted ontologies means that providers record important variables in different ways. Several interviewees commented that processes for sharing health-related data must easily fit into existing clinical systems, or providers will not adopt them. Data quality and the ability of researchers to understand the nuances of the recorded data were also described as critical aspects of data sharing (table 2). Finally, human error can impact data quality. For example, data has to be manually entered into many registries. This process takes considerable time (and therefore money) to enter and manually check the data.

Legal regulations are being interpreted differently

Interviewees were generally clear that the revised Swiss privacy laws and EU GDPR were similar and consistent across the principles and approaches, allowing most contracts to be fairly straightforward. Some felt the nuances were important and remained unsettled (table 2). However, many noted that if data are shared with partners from the UK and US, laws are quite different, impacting project-related data sharing. In particular, the US and UK are seen to define anonymity and pseudo-anonymity differently (both

Table 2:

Example quotes illustrating each data theme.

Theme	Description	Example quote
Logistics of Data Sharing	The actual data transfer is not a bottleneck.	I don't feel that the actual transfer of data is a real bottleneck. The only bottleneck … would be if we're talking about very large data sets, which are just too big to be sent in a couple of minutes. But also, that is not that much of a bottleneck in my experience. (Clinical trial centre team member)
	The logistics are time-consuming but have been eased through various Swiss template creation efforts.	… the templates are a good approach to standardising it, which is very important. It makes things much, much more easy than if we didn't have that. They have also incorporated a federated queries system for data, but I don't know how much this is used as we are not researchers, but we were involved in creating this and defining how it should work … and I guess things are handled well with the partners we have. I mean, they're collaborating well. And it's most important that everybody talks to everybody so that the flow of information is not stopped. (Data protection officer)
General Consent	The general consent is helpful, but some remaining uncertainties and challenges impede data sharing.	One is the general consent that is now with the hospitals, and that's a very useful thing. I think. Once we have these data that are consented, we can relatively freely work with that, and that's extremely useful, and I do not see major issues with that. Besides that, there might be different versions, and then it gets difficult to assign which patient agreed to what exactly. And if you have different versions, it's a huge mess. … The second thing is the informed consent. If you're doing a specific study, even with an intervention, you need a specific informed consent. And since we have … hundreds of studies now, it gets difficult because they [are] very specific. … So you have to look at it one by one to see what is allowed with this stuff and what's not so. The specificness of this informed consent makes it difficult to generalise. (Researcher)
Data Quality	There are challenges to collecting and interpreting data from clinical sources for research purposes.	Not only how the data are collected, but also the understanding of the data is not necessarily the same from institution to institution. … Researchers who don't have a clinical background, they don't understand the possible limitations of the data. Just give me the blood pressure. What's all the fuss about the blood pressure? But when you are a clinician, then you know that it's a huge difference if you are measuring it in the emergency department under such conditions or if it's an ambulatory visit or if you are lying down or if it is an intra-arterial blood measurement. And that's clear for us. And we have a feeling for how to deal with that. But the data scientists don't have this background much as they are just interested in the data. "I don't want to hear about the possible biases and limitations that might be attached." That's something that I fear. There should be more cooperation between clinicians and data scientists for more training. (Clinical trial centre team member)
Data Ownership	Participants were unclear about data ownership issues.	Well, in principle, those samples belong to the patient, and we are more like administrators of these samples. This whole ownership stuff is a relatively difficult thing because if I am employed, for example, with the [hospital], and I'm doing the clinical study and obtaining samples, what would be like a ranking of ownership, the patient of[sic] top. But then it's the hospital. Or is it me? Not really sure. And then, if it's a project funded by public institutions, then I would somehow feel obliged not to make obstacles for other researchers because there's a public interest maybe in samples. (Researcher)
		I think we get quite a nice access to the data by most institutions here in [University Hospital], but if you ask them directly, they would be very hesitant because there is still the mentality that the data primarily belongs to the institute or even to the physician who collected the data. But I think this is a bit outdated thinking, to be honest. … I believe that it's wrong to say that the data only belongs to the patient because the patient, without the help of our knowledge, our machines, our algorithms and so on and so forth, would never have reached this level of information. And the patient can often not even understand what it means. But I also think that data is not something that is just a number or just black or white data. It's also the knowledge, the contextual knowledge, what to do with a certain value, how to interpret them. I guess a pragmatic approach is to say that data is a collaboration between a patient and people who try to help the patient and some of the patients. (Researcher)
Legal Issues	Privacy laws.	I mean, it's basically one question, mainly. Is it applicable for Switzerland, or is it not? And if you ask five people, you will get six answers. … if you want to know my opinion, I think, and also my boss's … we think it was not made for the research context. (Lawyer)
	Anonymisation can seem legally straightforward but can be hard to operationalise in practice.	Now, from a legal perspective, this is quite simple. But I mean, [it] is just some words giving a phrase. And the phrase is "it's anonymous from a legal perspective" when you cannot, with reasonable measures, track it down to a person. Now, of course, that expression of reasonable measures is fluffy. [T]here are research papers from legal or from social studies telling you how far you can go. But there is no hard frontier of] what you can tell people. But of course, it's actually, I guess it's clear what coded means, [that] somewhere there's a key related to a person. But, they really had some difficulties in defining what have to do to make it anonymous. (Data security officer)
		They confuse what is encrypted and what is anonymised. And then they often say yes, they are anonymous to the outside world; the others don't know and don't see it. … There is no clarity that a key actually has to be destroyed for the data to actually be anonymised. And they also have no idea which technical solutions are available today to anonymise…. And I just think that I'm on the verge of genetic data … being so advanced that it's not anonymised anyway. (Ethics committee member)
		Yeah. I mean, you know, when I did the terms or the definitions, whether data is correctly anonymised or just pseudo-anonymised. And this is made clear in the Swiss Human Research Act. But if I see the data is anonymised, I mainly in all cases, I go back to the PI or to the study nurse or … coordinator asking, listen, is it really anonymised? … [It] is important because if you have really correctly anonymised data, the ethical committee is no longer competent … And so the law, the Human Research Act, isn't applicable anymore if it's really anonymised. (Lawyer)
		Well, in the beginning, it seems relatively straightforward, but the more you dove into the topic, the more complex the matter. … For example, if you want to have a demo data set for testing a tool or something, then it's not important that the values are very accurate. So you can blur that. That's how you can introduce jitter. … It can even be mixed between the patients. … It gets problematic when you need more information or you want to share the dataset in a more meaningful way. How big should the jitter be so as not to mess up your results? (Researcher)
		I would be extremely pleased if we had an instrument, a uniform instrument, so that we could then design anonymisation in the same way as other centres, for example. But the instrument that I just saw some time ago was much too extensive and too extensive for me. (Data security officer)
Culture of Data Sharing		[I]f you look on the timeline, the generating of data and include everything … writing a protocol, thinking about the methodology, getting the ethics approval, recruiting patients, which is challenging [sic] to find a patient to do all the testing, to have the query process to have the data at some point in the necessary exhaustive list and ready. This takes about 99% of the time. And 1% of the time is the actual analysing of the data. And actually, in data science, it's exactly the same. But people don't understand that. They think 1% is give [sic] me the data and 99% of the time is analysing it. But it's not true. The generation of the data is the real part of the project, and … people don't understand it. (Clinical trial coordinator)
		From an institutional viewpoint, the data belong to the institution, and we need to protect them because the researchers need to work on the data, and we need to keep them secure because there is evident competition within the academic sector who publishes first and then who owns the best data has better chances to publish. And this whole system is based on a peer publication type of competition. So, over the course of time, probably we created those barriers that made the protection of the data easy, but then the sharing [is] more complicated. (Bioinformatician)
		If you, for example, have a huge dataset that covers a whole set of patients with a specific condition, the hospital might not be very keen to send that to another institution. And we also have some, of course, competition between the different locations. And then you have to surrogate the agent with the [University]. And then it's always the question. If I have a collaboration with the [University], is it good if I send a huge data set of [hospital] patients to the [University], where they know they are very closely connected to [University]? And this could be a competitor …. That could be a problem. (Researcher)

Sustainability	We are building up a lot of new stuff. That's exactly the part where it's rather comparatively easy to get the funding. But you need infrastructures for data sharing, and they need to be kind of kept available over a longer period of time, and they need to be operated and maintained according to a high standard. … and it's like that's the psychological aspect. The researchers need to believe that they are available long term because it's clear that you are there. You can consider all the work you are putting into making data available. You can consider an investment, and then you want your sales, the return on the investment five years later, ten years later. But it means that the places where you put those datasets, we need to believe that they are still available in 5 to 10 years and operate to this high standard. And so, there is a kind of connection between the decisions of the funding organisations and what the researchers believe is going to happen. So you do not, on the one hand, ask the researchers to pay for such an infrastructure to be operated long-term. That doesn't work. No way, also not in Switzerland, so there needs to be continued funds to continue this after you have paid up this whole data sharing extraction. And then, because of the bad experience in the decades before, you need to get to the point where the research has been infinite, and that takes extra time. Even when the funding organisations have made up their mind, it still needs some time until it has sunk into the research as, yes, this is our infrastructure, and this is going to be available. If we are preparing something and putting it there, we are going to have it in five years. (Bioinformatician)

from each other and the GDPR), and the US laws have less privacy protection. Additionally, US contracts were seen as long and complicated, especially regarding any industry data sharing.

Data ownership was explicitly mentioned in about half the interviews. While this did not appear to be a bottleneck to data sharing, it was evidently a confusing area for the interviewees. There was no consistency regarding who interviewees perceived as "owning the data" (table 2). Several mentioned they felt the data belonged to the individual it originated from. However, others questioned whether it belonged to the physician who collected it, the institution, or another entity.

Every interviewee mentioned issues around anonymisation during the interviews. A few felt this was a straightforward issue. Others commented that it appears straightforward and binary at face value but is very complicated in practice. They described many levels of confusion and difficulty, including confusion between anonymised data, which cannot be linked back to research participants, pseudo-anonymised (coded) data, and encrypted data, which is transferred in a manner such that only authorised parties can access it (table 2). Interviewees also described the complexity of actually anonymising data in a way that maintains its research value. Some desired a tool that would uniformly help anonymise data so researchers and data security officers would feel confident in the anonymisation process. Lastly, several interviewees wondered whether data can ever really be anonymous in a small country like Switzerland and whether genetic data can ever be anonymised.

Creating a sustainable "culture of data sharing" in Switzerland

Many interviewees questioned how Switzerland could build a culture of data sharing and promote sustainable practices, such as long-term data curation and maintenance. Data sharing was described by several as needing to be seen "as a 'service', not a 'thing,'" and that data sharing should be considered as a more important goal of research than data protection (table 2). Several interviewees also described that health data sharing is moving toward focusing on data access rather than sharing. However, to make sustainable data access achievable, they stressed the need for the infrastructure to exist, be maintained, and be funded. Finally, regarding sustainability, interviewees stressed that if researchers and institutions are going to spend time curating their data to share it, they need to have an intrinsic drive to do so, seeing added value (table 2). They also stressed that there should be a system to acknowledge the intellectual contributions in sustaining research data and that funding should support sustainable data-sharing prac-

tices concerning data reuse from a technical- and governance-related level after the initial grants are completed. Finally, the inherent competition between researchers, universities, and hospitals (table 2) was seen as both a barrier and a potential way to create value scenarios in data sharing.

Discussion

Our study presented the results of interviews with 17 data-sharing expert stakeholders in Switzerland regarding the perceived bottlenecks within and outside of Switzerland. It found that, without focusing on its interoperability aspects, the actual data transfer process is not seen as a bottleneck and that logistics have improved with the development of harmonised templates (e.g. by the SPHN, SBP, and Swissethics) and the adoption of general consent across Switzerland. Among interviewees, there were differences of opinion and unclear areas remaining around legal issues, including how privacy laws in Switzerland compare internationally, the process of reconsent after general consent, and the process and definitions of anonymisation or pseudo-anonymisation. Barriers still exist regarding data quality, the use of routinely collected health data for research purposes, and the creation of a culture to support sustainable data sharing.

Concerns about the impact of inconsistent data quality and interoperability logistics have been well documented in the data-sharing literature. Many studies have focused on interoperability logistics, such as the need for consistent EHR adoption [11] and the development and use of specific ontologies [12, 13]. Our study demonstrated that Switzerland is seen as having made progress regarding many of the logistical barriers that challenge data sharing, such as developing harmonised templates and addressing general infrastructure issues at a national level [14]. However, the existing barriers go beyond the development and selection of ontologies; more work needs to address how to get data-providing institutions to implement existing semantic standards and get clinical providers to use ontologies proactively and in a way that minimally impacts the clinical process. Importantly, a procedural maze exists to allow data sharing to occur, and our interviewees felt that while it had become clearer in recent years, there remains a need for additional education and support to help facilitate the data-sharing process in Switzerland and elsewhere [12, 15].

As Vayena et al. [16] stated, "[m]ost of the debate about big data uses for health purposes has focused on privacy." Significant literature from Switzerland [4–6] and internationally [17,18] describes research participants' concerns about privacy and their desire that a clear purpose exists when data is shared. An expert stakeholder study in

Switzerland conducted in 2018–19 also demonstrated significant fears and frustrations by researchers regarding Swiss and international privacy laws [8]. As such, at the start of our study, we anticipated that the major bottleneck to data sharing would be confusion and misunderstanding about privacy laws and concerns about privacy in data sharing. Nonetheless, while our results show that some confusion remains, the general feeling from our interviewees was that privacy laws were not a major barrier to data sharing in 2021. There are several possible reasons for this difference over time. First, the implementation of general consent in Swiss hospitals (in 2017) and the publication of a harmonised version (in 2018) and its subsequent high acceptance [19] do seem to have made a difference for data sharing across Switzerland, despite the remaining confusion arising from institutional differences in the consent form. Second, the GDPR was enacted in 2018. In response, Switzerland has issued a revised Data Protection Act that comes into effect on 1 September 2023 [20, 21] and has had an adequacy decision for data protection since 2000 that is still in effect [22]. However, the EU is currently examining the revised Swiss Data Protection Act vis a vis the GDPR and is expected to issue a new decision. Our data suggests that the main legal uncertainties centre around how to define and operationalise anonymous data and data ownership issues. The "data ownership" issue is not unique to the Swiss context. It is evoked frequently in the international discourse in data access and sharing and often becomes a conversation stopper given the complexities surrounding both "data" and "ownership" in the digital space. Substantive questions have been raised about the usefulness and suitability of property frameworks for health data [23]. While we do not anticipate a resolution to this challenge, it is essential to acknowledge that it does not prevent further pursuit of ethical data uses. Several data governance models have been advanced that allow data access and sharing without necessarily using a property framework approach [16, 24].

Related to the issue of privacy is anonymisation. Our data showed that interviewees are still confused about the specific definitions of anonymisation and pseudo-anonymisation (sometimes called coded or de-identified data) and worry that the definitions vary internationally. There is also confusion between "encrypted" data, which is transferred such that it can only be accessed by authorised parties, and "coded" data; this confusion, which the SPHN Data De-identification Task Force has previously described in May 2022, may in part occur because the Swiss law refers to "verschlüsselte Daten" (encrypted data). Researchers in our study also expressed concerns about how to anonymise health data without making data significantly less useful for future research. Our study reinforces the confusion about these terms described by Chevier et al. [25] and suggests that health-related data sharing would benefit from clarifying definitions, perhaps along with examples of how to achieve the anonymisation steps. It may also be that new processes such as federated data sharing or approaches such as multiparty homomorphic encryption [26] will help address these issues.

Most importantly, interviewees emphasised two key concepts to move towards a sustainable process for data sharing: (i) promoting and supporting a culture of data sharing, and (ii) the need to transition our thinking process towards data access [27] rather than data transfer processes. Our study mirrored the findings of Geneviève et al. [7] in finding that the Swiss (and indeed international) academic environment created much pressure against data sharing by pushing researchers to publish data and research for the sake of "credit", with data sharing not valued in the same way. Many other studies assessing challenges in data sharing have also shown a need for academic and professional incentives in data sharing [15, 28–31]. Another aspect common to our study and the literature is the need for cost-sharing around creating sustainable data-sharing infrastructure [28].

Finally, regarding changing towards a focus on data access rather than data sharing, there are notable developments in the approaches used by large health data organisations, such as in the various National Health Service organisations across the UK [32]. Enabling data access within trusted environments is a new policy aiming at enabling trustworthy health data uses by avoiding risks associated with actual sharing and moving data amongst different entities. Such an approach is conceptually and practically different from typical data-sharing models.

Study limitations

Our study was limited in that it addressed a small number of individuals in Switzerland who perform various data-sharing tasks and functions, with an explicit focus on researchers involved in the Swiss Personalised Health Network data ecosystem. This ecosystem has been conceived as a model environment to facilitate data access and sharing and, therefore, provides an excellent opportunity for studying bottlenecks given the effort put into improving data flows. In particular, our participants only included two research scientists who share and receive data. Our study was also limited by a single researcher performing the interviews and coding processes. However, we took steps to ensure that the entire research team was involved in the theme selection and verification process. Finally, as with all qualitative research, the reader should not attempt to generalise the potential prevalence of attitudes presented in this article.

Potential areas for action

Given our study results, how can Switzerland continue to improve its national data sharing and access capabilities? First, given the many areas with questions and uncertainties, a common understanding of the requirements and close coordination and harmonisation of processes, practices, and interpretation of the law is needed at the national level. While streamlined significantly in recent years, logistics could still benefit from additional standardisation and transparent communication, including the potential consideration of a single contact approach at each institution, templates recognised across Switzerland's healthcare institutions and regulatory bodies, and a nationally harmonised general consent process. The guidance could define what it means to anonymise and pseudo-anonymise health data and how this differs legally across different countries, including case examples. As processes become better defined, more education for researchers and individuals at all levels of the data-sharing enterprise

would be helpful. Investigators should be encouraged to work with the relevant data governance and clinical trials teams within their institutions, who have defined the required processes and have knowledge and experience in streamlining them. Additionally, organisations could create FAQ documents about how various international privacy and data-sharing laws compare to Swiss laws or clarify some of the questions around the general consent process. Finally, future studies could identify a broader population of such experts across Switzerland's entire health data ecosystem and consider a quantitative approach to determine which aspects in our study are broadly endorsed and by whom.

Data sharing statement

The authors will share de-identified transcripts with participant permission or sections of coded data upon reasonable request to the corresponding author.

References

1. Institute of Medicine (US). The Benefits of Data Sharing. US: National Academies Press; 2013.
2. Coman Schmid D, Crameri K, Oesterle S, Rinn B, Sengstag T, Stockinger H; BioMedIT network team. SPHN - the BioMedIT network: A secure IT platform for research with sensitive human data. Stud Health Technol Inform. 2020 Jun;270:1170–4.
3. ELSI projects [Internet]. SPHN. SPHN - Swiss Personalized Health Network; 2019 [cited 2023 Sep 16]. Available from: https://sphn.ch/network/projects-old/elsi-projects/
4. Brall C, Berlin C, Zwahlen M, Ormond KE, Egger M, Vayena E. Public willingness to participate in personalized health research and biobanking: A large-scale Swiss survey. PLoS One. 2021 Apr;16(4):e0249141. http://dx.doi.org/10.1371/journal.pone.0249141.
5. Brall C, Berlin C, Zwahlen M, Vayena E, Egger M, Ormond KE. Public preferences towards data management and governance in Swiss biobanks: results from a nationwide survey. BMJ Open. 2022 Aug;12(8):e060844. http://dx.doi.org/10.1136/bmjopen-2022-060844.
6. Pletscher F, Mändli Lerch K, Glinz D. Willingness to share anonymised routinely collected clinical health data in Switzerland: a cross-sectional survey. Swiss Med Wkly. 2022 Jun;152(2324):w30182. http://dx.doi.org/10.4414/SMW.2022.w30182.
7. Geneviève LD, Martani A, Perneger T, Wangmo T, Elger BS. Systemic Fairness for Sharing Health Data: Perspectives From Swiss Stakeholders. Front Public Health. 2021 May;9:669463. http://dx.doi.org/10.3389/fpubh.2021.669463.
8. Geneviève LD, Martani A, Elger BS, Wangmo T. Individual notions of fair data sharing from the perspectives of Swiss stakeholders. BMC Health Serv Res. 2021 Sep;21(1):1007. http://dx.doi.org/10.1186/s12913-021-06906-2.
9. Creswell JW, Poth CN. Qualitative Inquiry and Research Design: Choosing Among Five Approaches. SAGE Publications; 2016. 488 pp.
10. Grady MP. Qualitative and Action Research: A Practitioner Handbook. Phi Delta Kappa International; 1998. 55 pp.
11. Auffray C, Balling R, Barroso I, Bencze L, Benson M, Bergeron J, et al. Making sense of big data in health research: towards an EU action plan. Genome Med. 2016 Jun;8(1):71. http://dx.doi.org/10.1186/s13073-016-0323-y.
12. Tacconelli E, Gorska A, Carrara E, Davis RJ, Bonten M, Friedrich AW, et al. Challenges of data sharing in European Covid-19 projects: A learning opportunity for advancing pandemic preparedness and response. Lancet Reg Health Eur. 2022 Oct;21:100467. http://dx.doi.org/10.1016/j.lanepe.2022.100467.
13. Gaudet-Blavignac C, Raisaro JL, Touré V, Österle S, Crameri K, Lovis C. A National, Semantic-Driven, Three-Pillar Strategy to Enable Health Data Secondary Usage Interoperability for Research Within the Swiss Personalized Health Network: methodological Study. JMIR Med Inform. 2021 Jun;9(6):e27591. http://dx.doi.org/10.2196/27591.
14. The SPHN Semantic Interoperability Framework [Internet]. SPHN. SPHN - Swiss Personalized Health Network; 2020 [cited 2023 Sep 16]. Available from: https://sphn.ch/network/data-coordination-center/the-sphn-semantic-interoperability-framework/
15. Anger M, Wendelborn C, Winkler EC, Schickhardt C. Neither carrots nor sticks? Challenges surrounding data sharing from the perspective of research funding agencies-A qualitative expert interview study. PLoS One. 2022 Sep;17(9):e0273259. http://dx.doi.org/10.1371/journal.pone.0273259.
16. Vayena E, Blasimme A. Health Research with Big Data: Time for Systemic Oversight. J Law Med Ethics. 2018 Mar;46(1):119–29. http://dx.doi.org/10.1177/1073110518766026.
17. Lynch F, Meng Y, Best S, Goranitis I, Savulescu J, Gyngell C, et al. Australian public perspectives on genomic data storage and sharing: Benefits, concerns and access preferences. Eur J Med Genet. 2023 Jan;66(1):104676. http://dx.doi.org/10.1016/j.ejmg.2022.104676.
18. Middleton A, Milne R, Howard H, Niemiec E, Robarts L, Critchley C, et al.; Participant Values Work Stream of the Global Alliance for Genomics and Health. Members of the public in the USA, UK, Canada and Australia expressing genetic exceptionalism say they are more willing to donate genomic data. Eur J Hum Genet. 2020 Apr;28(4):424–34. http://dx.doi.org/10.1038/s41431-019-0550-y.
19. Griessbach A, Bauer A, Jörger Lebet F, Grossmann R. The concept of General Consent in Switzerland and the implementation at the University Hospital Zurich, a cross-sectional study. Swiss Med Wkly. 2022 Apr;152(1516):w30159. http://dx.doi.org/10.4414/SMW.2022.w30159.
20. Portal SM. New federal act on data protection (nFADP) [Internet]. [cited 2023 Sep 16]. Available from: https://www.kmu.admin.ch/kmu/en/home/facts-and-trends/digitization/data-protection/new-federal-act-on-data-protection-nfadp.html
21. Privacy Law [Internet]. [cited 2023 Sep 16]. Available from: https://swissprivacy.law/wp-content/uploads/2021/02/20210211-Tableau-comparatif-nLPD-et-RGPD.pdf
22. Federal Data Protection, Information Commissioner (FDPIC). FDPIC Adequacy Decision [Internet]. [cited 2023 Sep 16]. Available from: https://www.edoeb.admin.ch/edoeb/en/home/datenschutz/international/angemessenheit.html
23. Liddell K, Simon DA, Lucassen A. Patient data ownership: who owns your health? J Law Biosci. 2021 Oct;8(2):lsab023. http://dx.doi.org/10.1093/jlb/lsab023.
24. Xafis V, Labude MK. Openness in Big Data and Data Repositories: The Application of an Ethics Framework for Big Data in Health and Research. Asian Bioeth Rev. 2019 Oct;11(3):255–73. http://dx.doi.org/10.1007/s41649-019-00097-z.
25. Chevrier R, Foufi V, Gaudet-Blavignac C, Robert A, Lovis C. Use and Understanding of Anonymization and De-Identification in the Biomedical Literature: scoping Review. J Med Internet Res. 2019 May;21(5):e13484. http://dx.doi.org/10.2196/13484.
26. Scheibner J, Raisaro JL, Troncoso-Pastoriza JR, Ienca M, Fellay J, Vayena E, et al. Revolutionizing Medical Data Sharing Using Advanced Privacy-Enhancing Technologies: Technical, Legal, and Ethical Synthesis. J Med Internet Res. 2021 Feb;23(2):e25120. http://dx.doi.org/10.2196/25120.
27. Access to Health Data [Internet]. HDR UK. 2020 [cited 2023 Jan 13]. Available from: https://www.hdruk.org/access-to-health-data/
28. Gudi N, Kamath P, Chakraborty T, Jacob AG, Parsekar SS, Sarbadhikari SN, et al. Regulatory Frameworks for Clinical Trial Data Sharing: scoping Review. J Med Internet Res. 2022 May;24(5):e33591. http://dx.doi.org/10.2196/33591.
29. Tenopir C, Dalton ED, Allard S, Frame M, Pjesivac I, Birch B, et al. Changes in Data Sharing and Data Reuse Practices and Perceptions among Scientists Worldwide. PLoS One. 2015 Aug;10(8):e0134826. http://dx.doi.org/10.1371/journal.pone.0134826.
30. Popkin G. Data sharing and how it can benefit your scientific career. Nature. 2019 May;569(7756):445–7. http://dx.doi.org/10.1038/d41586-019-01506-x.

31. Bierer BE, Crosas M, Pierce HH. Data Authorship as an Incentive to Data Sharing. N Engl J Med. 2017 Jul;377(4):402. http://dx.doi.org/10.1056/NEJMc1707245.

32. Joining up the dots: driving innovation, research and planning through Trusted Research Environments [Internet]. NHS Transformation Directorate. [cited 2023 Jan 13]. Available from: https://transform.england.nhs.uk/blogs/joining-up-the-dots-driving-innovation-research-and-planning-through-trusted-research-environments/

33. QSR International Pty Ltd. (2020) NVivo (released in March 2020), https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home

Appendix: Interview guide

The interview guide is available for download as a separate
file at https://doi.org/10.57187/s.3538.