Journal of Library and Information Studies 19:2 (December 2021) pp.25-41
doi: 10.6182/jlis.202112_19(2).025

Patient Data Accessibility for Biotech and Medicine Industry Start-ups in Taiwan

Wei Jeng[1], Zih-Han Wang[2], Wayland Chang[3]

Abstract

In Taiwan's current biotechnology and medicine industry (BMI) landscape, there exists a data accessibility-ethical dilemma that might lead to an overall health data accessibility and availability supply-demand gap. Focusing on Taiwanese BMI startups' perceived data accessibility, this article—with a qualitative approach—altogether analyzed such startups' current challenges and circumstances. We conducted semi-structured interviews with 17 participants as different stakeholders, including 4 physicians, 5 BMI startup workers, and 8 with both roles. Also analyzed were several actual data accessibility issues that Taiwan's BMI startup community often encounters during data acquisition. Ultimately, this study confirms a rather inconvenient truth that data acquisition costs are rather very high – not only in monetary-wise, but also perceived efforts-wise in response to Taiwan's rather rigid biotechnology regulatory regime. Also found were that many of the interviewed BMI startups had voiced hopes for more regulatory transparency plus well-roundedness than those of more regulatory relaxations.

With an interview approach, we found still more rooms of improvement in the realms of the startup itself, BMI startup accelerators, and government in Taiwan. Thus, this study has three recommendations: First, better startup's data literacy to better respond to relevant data requests. Second, better governmental transparency and well-informed regulations. Third, about the aforesaid complex regulations, better mentorship by such startup accelerators. Also observed were startups' various proactive attempts to strike a strong ethics-development growth ethics. Thus, this study provides deeper insights into building more mutually beneficial approaches in the data accessibility issues realm for Taiwanese BMI startup communities' different stakeholders. Namely as such startups are often at the crossroads of burgeoning expansionary visions and sometimes stifling regulatory regimes.

Keywords: Health Data; Data Accessibility and Reusability; Data Infrastructure; Biotech and Medicine Industry Startups

1. Introduction

Technological advances altogether allow people to collect, disseminate, process, and store scientific data on a greater scale. Yet, when data involve human beings, immediately arisen are ethical concerns about data privacy and sensitivity.

Nonetheless, proper health data use can benefit medical innovations. With proper personal health data analyses, highly possible are lessened prescription errors, greater potential disease detections, and even pandemic (e.g., COVID-19) preventions. And with such use, more modern techniques such as data mining and machine

[1,3]Department of Library and Information Science, National Taiwan University, Taipei, Taiwan
[2] Information School, University of Washington, Seattle, USA
* Corresponding Author: Wei Jeng, E-mail: wjeng@ntu.edu.tw

learning can altogether enhance the medical sciences (Bindley, 2019). Yet at times, personal benefits do not always outweigh people's potential privacy concerns. As health data misuse might lead to dire consequences such as identity fraud.

With the aforesaid possible issues, this study analyzes the current data accessibility landscape in Taiwan's biotechnology and medicine industry (BMI). Overall, greater health data access allows Taiwanese BMIs to effectively improve their respective prototypes, product marketing, and raise more funds. Yet, it is unclear to what extent the Taiwanese BMI community and employees are authorized to release, access, along with apply health data during the clinical testing and product development stages.

Currently, the most notable challenge for Taiwanese BMI startups to access patient data is the tension between innovation and the strictness of the application process. A BMI startup needs to go through several regulatory approval hurdles in obtaining data/patients and market such startups' products. Altogether, these hurdles are financially costly and time-consuming and can be a hinder for a startup's growth. About Taiwan's biotechnology regulatory regime, any necessary approvals are given by Taiwan's Food and Drug Administration (FDA). Individual institutions-wise, such approvals are per such institution's respective Institutional Review Board (IRB) unit. Nonetheless, achieving both FDA and IRB approvals needs significant efforts. With IRB as an example—as part of the review process—Taiwanese BMI startups often need to have physicians serve as the project's principal investigators. Such a need is often a startup's Achilles heel as very few startups even have existing physician contacts at the onset. And upon a months-long IRB review process when such startups eventually achieve approvals, startups thus often undergo a costly health data acquisition process (e.g., apply for patient records, recruit patients to independently collect data).

Various countries worldwide respectively host a national health database. Examples include Taiwan's National Health Insurance Research Database (NHIRD) and the United Kingdom's Clinical Practice Research Datalink (CPRD). Yet and regardless of jurisdiction, even if a BMI startup company were to overcome the aforesaid costly procedures, the national health database (e.g., Taiwan's NHIRD) often operates as a data enclave. As data consumers need to both submit preferred variables beforehand and physically access the database under the watch of such database's staff—process often deemed inconvenient for BMI startups.

The notion of BMI stakeholders in this study comprises BMI startup company employees and medical doctors (hereafter: physicians). To analyze the current states of Taiwanese health startups' perceived data accessibility and challenges, this study was qualitatively approached with interviews of seventeen persons involved in Taiwanese health-related startups and/or work as physicians.

This study focuses on a trio of research questions:

RQ1. What are Taiwanese BMI startups' data needs? Such needs occur per what occasions?

RQ2. What health data access issues do such startups face in the project developmental stages?

RQ3. What are the ideal approaches that such startups perceive to access data that they need?

Taiwan's bio-medicine industry has steadily progressed. First, as of July 2020, Taiwanese BMIs had a total revenue of 559.7 billion NTD (roughly 20.1 billion USD). Second, Taiwan is often ranked in the top tier of top-performing jurisdictions in health care (CEOWORLD Magazine, 2021; Economist Intelligence Unit, 2017). Third, per both the "New York Times" and "Global Open Data," Taiwan is often ranked in the top tier of government data openness (Carroll & Frakt, 2017; Open Knowledge Network, n.d.). With such successes in mind, Taiwanese BMIs' data accessibility case—coupled with its possible challenges—can serve as a vital reference for other jurisdictions to consider.

2. Literature Review

There has been a lack of literature focusing on how BMI startups access the patient data in hospitals, nor now they establish their own approaches to collect patient data. More recent literature has nevertheless centered on the accessibility of clinical trial data, including patient data, clinical trial records (e.g., Green et al., 2015; Sudlow, Branson, Friede, Morgan, & Whately-Smith, 2016). This section raises startups' data accessibility issues by reviewing tensions between data protection and technology innovation and reviewing the regulations in Taiwan.

2.1 Tensions between data protection and technology innovation

Tensions between data accessibility and technology innovation exist and become more notable these years. Martin, Matt, Niebel, and Blind (2019) investigated how data protection regulation affects technical innovation among EU startups during the effect of General Data Protection Regulation (GDPR). Their finding suggests that the GDPR on startup innovation can be complex: it instantaneously drives and constrains the innovation at-hand (Martin et al., 2019). Boyacioglu and Yıldız (2021) also arouse the problem within data accessibility and startups: it can be difficult to reach the best solution that can ease such tensions between strict data protection and unfair competition.

While data accessibility can be understood as being able to obtain data whenever needed (Strong, Lee, & Wang, 1997), in Taiwan, when a BMI startup needs to access patient data to develop products, two databases are often applied to: First, the Hospital Information System (HIS) that are both built and managed by the hospitals themselves. Second, the NHIRD is managed by the Taiwanese Ministry of Health and Welfare's National Health Insurance Administration. Both such databases' accesses require researchers to undergo the aforesaid IRB of each database's governing authority.

Compared to that of the NHIRD, the HIS' information is recorded by physicians when diagnosing patients, and such information is considered firsthand data. Namely as the HIS contains both a patient's detailed personal information and medical records. When the HIS is released to researchers, the relevant Information Technology Offices then anonymize the patients' respective identities. For both a safe and sustainable HIS environment, existing anonymization mechanisms have nevertheless been far from perfect. As more technologies both enable connection and integration across different databases, researchers have increasingly

been able to speculate about patient identity. An investigation on a certain Taiwanese hospital's database has surprisingly shown that with just the patient's birthday information, up to 3,354 of 190,027 (roughly 1.8%) patient identities can be recognized (Chuang & Tsao, 2012). Also, knowing which health specialty the patient has attended to further helps de-anonymize the dataset: 131,168 patients can thus be recognized (69% of the database population).

From some views, the NHIRD is a second-hand database that contains information meant for healthcare reimbursement. As hospitals often collect doctors' advice in applying for reimbursement from relevant authorities in compliance with relevant Taiwanese insurance laws. Such practices have aroused serious debates about possible citizen privacy violations. Namely as most people are often unaware that their respective information might be altogether applied for research purposes. For instance: per Chang (2016) and in a ruling struck in 2014, the Taiwan Taipei High Administrative Court Judgement No. 102-SU-36 (Tsai v. NHIA) ultimately ruled that Taiwan's National Health Insurance Administration had the authority to collect—without consent—data subjects' personal health information for the Administration's goal of promoting public welfare. This Court's key reasoning was that such "data transfer was permissible because it was conducted for academic research in the public interest."

With the above in mind, patient recruiting is another practice highly related to patients' own data collection. Startups often recruit patients to record firsthand data by themselves. Hospitals, convalescent centers, friends, and families altogether are all resources for startups to find their respective patients. Regardless of the channel that startups choose, startups are compelled to actively recruit patients – never an easy task. Such practices have likewise stoked some ethical controversies.

At the onset, taking advantage of patients' predicaments is clearly an ethical controversy. Patient recruiting process-focused research has shown that some startups are taking advantages of patients' desperation about their respective diseases or their socioeconomic status to attract them to taking part in the tests (Patra & Sleeboom-Faulkner, 2009). For example, a startup focused on stem cell therapy targeted patients with untreatable diseases or those who have already spent a large amount of money on standard treatment, along with to avoid as much responsibility as possible.

2.2 Data accessibility process and regulations in Taiwan

While health data at a national scale has rapidly increased, there have also been increased concerns about patients' possible privacy and ethical issues. Overall, Taiwan's FDA and IRB premises mostly reflect the USA's. Especially as persons with serious diseases are sometimes treated as an underprivileged group (i.e., economically disadvantaged, lacks physical autonomy), such group almost never gains as much information about a product's risks as that of medical personnel or startups. With this inequity, while one argument might voice that health data increases boost both a health information system's improvements and openness, an opposing argument might voice that a patient's privacy should be of utmost priority.

Before startups market and sell their respective products, such startups need to both achieve relevant FDA approval and pass the FDA's examination. Thus, startups first have to provide data to prove efficacy that in turn requires IRB approval.

The U.S. Food and Drug Administration's (FDA) mission is to altogether protect people by ensuring that foods, drugs, and medical devices are both safe and efficacious. Namely as the FDA is responsible for supervising drugs, biological products, medical devices, animal drugs, and food addictiveness. Only upon products undergoing the FDA's pre-marketing approvals can such products be both marketed and sold in the United States. Comparatively and like the FDA's, Taiwan's Food and Drug Administration (TFDA) holds similar roles. Besides managing both drugs and foods that are to be subsequently sold in the local Taiwanese market, the TFDA is also responsible for food and drug risk assessments.

Regardless of jurisdiction, each organization—those that offer patients' data or arrange startups' patient recruitment—has its own IRB unit that examines if the startup project and/or data usage at-hand conforms to relevant ethical values. As the IRB focuses on a justice-like mantra for patients, anything involving human subjects needs prior IRB approval so that the at-hand project's overall ethical integrity is upheld.

And regardless of jurisdiction, before applying for an FDA license, a startup altogether needs to undergo the aforesaid IRB, complete any relevant clinical trials upon approval, and obtain empirical data. Also, a mandated FDA application document includes that of testing results. Thus, both the FDA and IRB are indispensable for startups in wanting their respective medical products to be both marketed and sold.

3. Research Design

This study aims to understand the multiple stakeholders' situations, attitudes, and expectations surrounding the BMI startups' ecosystem. To answer this study's aforesaid research questions, conducted were seventeen in-depth interviews encompassing different stakeholder groups. Such interview approaches were applied for two reasons: First, this study wanted to directly communicate with stakeholders and capture their respective insights in certain contexts with expectations that the interviewees have enough time if they find certain topics difficult to elaborate. Also, proactively needed was to observe each interviewee's reactions so that possibly this study's interview tone plus expression were properly modified to ensure that all the asked questions were done so in both neutral and inoffensive ways (Mack, Woodsong, MacQueen, Guest, & Namey, 2005). Second, as each BMI startup's product varies and has differing developing processes, anticipated were flexibility needs to appropriately needing adjust or change the sequence of questions to get deeper understanding in words that we are not familiar with. As detailed more below, a somewhat semi-structural design was applied.

As for the participants, to better understand BMI startups' situations, attitudes, and expectations, individuals involved in the BMI startups themselves are our target population (Group A). For Group A, we focus on gaining an overview of healthcare startups' needs and experiences. As mentioned before, the role of physicians within the data collection process is very crucial. We decided to include physicians who are currently working in a hospital in the

sample (Group B). For Group B, we anticipate to understand medical personnel's attitudes about the current circumstances. Finally, we consider individuals who can be classified into both immediate groups mentioned above to observe physicians involved in their own startup business and how their respective identities can help or constraint their data collection tasks; we especially include them in our sample as Group C.

Among Group A's five participants we interviewed, three of them worked on BMI startups and had no in-hospital experience in the past. This group reflected one of this study's goals in analyzing BMI startup stakeholders who had the fewest medical resources at-hand, coupled with an overview of healthcare startups' needs and experiences. Group B consisted of in-hospital physicians. Another of this study's goals is to understand medical personnel's current data accessibility attitudes plus circumstances. As medical personnel are often at the nexus between patients and startups/hospitals and patients – thus holding a vital role in the data access and collecting process. Lastly, Group C is made up of medical personnel who held prior startup experience. As upon facing data-accessing difficulties, medical personnel might have different opinions toward data accessing and collecting issues. Namely as another of this study's goals is to discern any possible differences from the answers given by both Group A and B.

3.1 Piloting and refining interview protocol

Piloting was applied to first assess if the interview protocol was properly designed. Upon testing with three pilot participants, this study's interview protocols were adjusted to make the interview questions more neutral and concise, along with to remind participants that they have the right to request that the recorder be turned off at any point. Table 1 details both the process and interview content for both Groups A and B. Group C's interview protocol is both combined and modified from Group A and Group B's. Why? Group C's interview was applied to both compare the attitudes and results with those of the other groups.

3.2 Case study of a BMI startup accelerator program

To gain an overview of the current situation, conducted was a case study based on a startup accelerator that was cooperating as an Industry-Cooperation Research Project with the National Taiwan University. The startup accelerator can be considered a platform in which different BMI startups stakeholders can be reached. For Group B, besides the startup accelerator's network, we reached out to physicians employed in local hospitals. One participant (C05) was originally placed in Group B. Yet, upon the interview's commencement, found out was that C05 had previously been in a startup project. Thus, the protocol was subsequently changed and C05 was re-arranged to become a Group C participant. The numbers of participants, their respective backgrounds, and work situations are altogether presented in both Figure 1 and Table 2.

After the interview, all the audio files were respectively transcribed into text files. The research team conducted an open coding process from transcription and captured needed information. Three different themes were rippled during the coding process: 1) occasion for data and patient access, 2) obstacles for startups to access and collect data, and 3) desired infrastructure

Table 1. Interview Protocols for Group A, B, and C

Stages	Description for group A (startups)	Description for group B (physicians)	Description for group C (mixed)
1. Warm-up.	I. Researchers introduce the study; acquire consent.		
	The interviewee describes his/her startup product, startup members' background, and current development stages.	The interviewees describe their respective hospital division.	The interviewees describe their respective hospital divisions, his/her startup product, startup members' background, and current development stages.
2. Understand data access and collection needs (Group A) or requests (Group B).	Interviewees are asked to describe all development stages and then explain at which stage the demand of patient data is highest. Also, the interviewees are asked about the impact of failing to access data.	Interviewees are asked if they have received requests for both data access and data collection.	Interviewees are asked to describe all development stages and then explain at which stage the demand of patient data is highest. Also, the interviewees are asked about the impact of failing to access data. On the other hand, interviewees are asked if they have received outside requests for both data access and data collection as a physician.
3. Practices to both access and collect data.	Interviewees are asked to elaborate the possible ways to both access and collect patient data. Also, asked were which channel was chosen and what the obstacles were.	Interviewees are asked to elaborate on the possible ways of accessing or collecting patient data in the general settings.	Interviewees are asked to elaborate the possible ways to both access and collect patient data as a combination identity.
4. Figure out startup side's expectations (Group A); Figure out medical personnel's attitudes about both data access and collection (Group B).	Interviewees are asked to describe the ideal ways to access potential users and collect patients' data. Also asked if other startup companies had tried this method and what the results are.	Asked were the interviewees' opinions on both the current situation and regulations, along with how he/she expects other medical personnel to react when facing data accessing and collecting needs from startup companies.	Asked were the interviewees' comprehensive opinions on both the current situation and regulations regarding data accessibility for BMI start-up communities as a combination identity.
5. Debrief.	Researchers expect to receive both responses and suggestions about the study from the interviewees.		

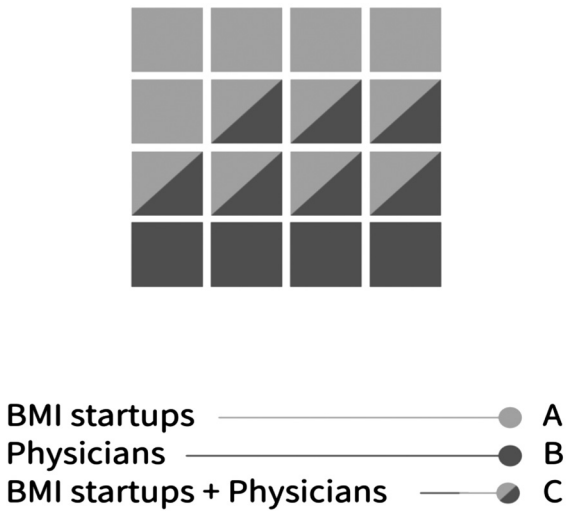

Figure 1. Participant Groups

of accessing data for BMI startups in Taiwan. The researchers' notes for open coding and the transcribed text yielded 14,433 words (21 pages).

4. Results

4.1 Occasion for data and patient access

In examining which stage of the BMI startup's development cycle had the highest data demand, found was that it was dependent on the product type. Per this study's participants, the innovative product development, before its public release, can thus be altogether divided into several phases: idea forming, prototype producing (i.e., testing), and the service or product marketing stage.

For the *idea forming* stage, as with the research lifecycle's idea forming, BMI startups at this stage need the patient's data to both confirm what the unmet needs are and to determine the potential market's size.

At the *prototype producing* stage, data at the prototype producing stage can be useful so that the startups can determine if the patient is simply the right person or which patients should be excluded from the prototype tryout. Both participants C02 and C03 worked on a project that produces oral fluid absorbing facial masks. As both C02 and C03 want to help post-stroke affected patients who had oral swabbing issues. Without reaching out to the patient and observing how the product works, the prototype would not be as reflected: "*There is a moment that we felt we need patients more, the prototyping phase. We need to know whether we should go for disposable or replaceable pads for our masks, that was a must (for recruiting patients)* (C02)."

At the *service or product marketing* stage, marketing is a stage that can be improved greatly with patient data that can altogether show which areas the patients that have the greatest needs. Such areas might involve the lack of relevant awareness about a certain disease or which patients are more willing to go through treatment.

Table 2. Overview of Pilot Participants (*n* = 3: A00, B00, and C00) and Participants (*n* = 14)

Groups	Participants	Gender	Age	Situation/occupation description
Group A (startups)	A00 Pilot#1	M	31-40	The pilot participant constructs a neater and more integrated infusion system in hospitals, and works as a regulatory affairs specialist in the team.
	A01	M	41-50	The participant works on a medical device preventing Obstructive Sleep Apnea, and the team members are mostly electronics engineers and mechanical engineers.
	A02	F	41-50	The participant develops a software that analyses patient data. The members of the team are mostly data analysts. The participant was a nutritionist before joining the current job.
	A03	F	21-30	The participant designs an app recording details of women's menstrual cycle. The members of the team are mostly designers and software engineers.
	A04	M	41-50	The participant develops a system that controls the risk under radiation therapy. The team members are mostly engineers.
Group B (physicians)	B00 Pilot#2	M	21-30	The participant is a senior (fourth-year) medical school student.
	B01	M	41-50	The participant is an emergency room doctor and had served on an IRB committee in a private, teaching hospital.
	B02	M	21-30	The participant is a resident doctor working in the Division of Radiation Therapy
	B03	M	41-50	The participant is a doctor working in the Division of Plastic Surgery.
Group C (startup-workers who are also physicians)	C00 Pilot#3	M	21-30	The participant is a senior level healthcare specialist working in a healthcare startup accelerator.
	C01	M	51-60	The participant is a nephrologist who runs a startup team focused on improving hemodialysis machines.
	C02/C03	M/F	21-30/ 21-30	Participant C02 is a physical therapist who designs face masks.
	C04	M	31-40	The participant was an emergency room (ER) doctor who had developed a system that analyses patient data.
	C05	M	31-40	The participant is a medical laboratory scientist developing a solution to improve therapeutic efficacy and control the side effects.
	C06	M	21-30	The participant is a physical therapist who designs assistive technology devices.
	C07	M	31-40	The participant was an emergency room (ER) doctor who develops a software system for decision-making support in patient care.

Whereas, patient data analysis-focused startups often have a mantra that "data is required in the product's whole developing lifetime." Namely as the approaches startups use to both access and collect data can sometimes vary much from the in-hospital channel to the self-outreaching channel.

4.1.1 Collaborating with physicians

FDA approval-wise, BMI startups in this research's sample are needed to go through the IRB to access or collect patients' data. In that case, collaboration with physicians is thus a necessary move for Group A. Yet, not everyone has the necessary social network to do so. And for persons who do not have physician friends and/or relatives, some startup participants chose not to spend so much effort to go through the IRB process (e.g., giving up in attempting to reach hospital patients). To counter this inconvenience, physicians are interviewed instead as reflected here: "*In fact, it is still necessary to interview doctors and medical staff in medical institutions... According to their experience to understand the background and condition of patients* (A01)."

4.1.2 Reaching clients at social media

We also found that many current startups apply social media. Four participants (A02, A03, C02/C03) have previously attempted to recruit both interviewees and users from Facebook. For follow-ups, both the LINE message app and/or Facebook Messenger have been used. The products team often uses social media to reach potential clients who often use—instead of being more interested in invasive-related treatments—mobile health apps and/or wear wearable devices. And to attract the general public, such companies often offer complimentary services and/or products. For example, as compensation, A03's company provides complimentary thermometers for its clients who help report a woman's basal body temperature (BBT) to enhance such woman's menstrual cycle prediction.

4.1.3 Reaching clients via personal network

Besides social media, startup personnel also turn to their respective real-world personal networks for potential help. Thus, for piloting, both friends and families have evolved as the first choice. In Group C, a startup even recruited a team member's patients to try the startup's (non-invasive) products. And a medical personnel had indicated that the key to successful patient access is the trust that has been gradually built with patients. Also, a successful startup pitch is often vital for patients in belief in both the product's efficacy and security.

Based on our interview results, some startups have asked Taiwan's Academia Sinica (which serves as Taiwan's national academy) and/or insurance companies for help. Yet, such startups shortly found out that such data resources are often of in-hospital so such organizations have no right to release any information. Other startups have attempted to visit patients or elders by randomly knocking on the doors of houses or rehabilitation centers. However, most of such requests have not been accepted.

Besides the methods startups that have used, some interviewees offered other possible channels. In Group B, a physician said that some private health examination centers have sold personal data with the patients' consent. Yet, few data about sensitive diseases are recorded in such databases. And some companies choose to build partnerships with health evaluation organizations, sharing the patients' data while the companies altogether offer

techniques, systems, and data collecting research assistants. Nevertheless, it is both expensive and often not an affordable resource for startups.

4.2 Obstacles for startups to access and collect data

When asked about health data access' main obstacles, the interviewees both shared their respective experiences and provided concrete examples that can perhaps be categorized into three aspects:

4.2.1 Physicians are hard to reach

In short, finding physicians for project assistance is a major obstacle. For example: "*If we don't have doctors in our team, it must be very difficult for us to acquire approval from IRB* (A04)." Per the interview results, startups often found it hard to acquire IRB approval if the project looked lucrative but without an academic or research purpose. Thus, startups must develop such a project into a research study before the IRB approval application. And startups often need to ask physicians to become such a project's principal investigator (PI). This is so that the IRB believes in the such project's academic value in medicine.

Another example: "*When a doctor spends 20 seconds explaining the patient-recruitment project to each patient, he/she has to work 2 extra hours every day* (B01)." A Taiwanese medical professional had grievances about hospital workload and even compared such work as "sweat factories." Thus, when physicians cannot even properly manage their clinical work, it is even more unlikely that such physicians will even offer to help startups. Also, a product's possible unknown risks often discourage physicians to help. As reflected by a Group B physician: "*Why would a doctor take risk of harming the patients just to become a PI without real power in a project?*" Besides such possible unknown risks, there are not many incentives at the onset for physicians to partake in startup products. And often the startup project somewhat operates like those of hospital research studies – thus the physicians cannot receive much benefits. With such circumstances, even if startups show their respective success potentials, often very few physicians are willing to participate in such aforesaid projects at the onset. Thus, it is rather difficult for startups at the onset to commence the IRB's examinations process.

4.2.2 The overhead of accessing and collecting data is very high

Each database has its own respective overhead costs and most startups will not purchase databases from a single hospital. Thus, startups are charged much if they are dedicated to obtaining both integrated and sufficient data from different organizations. If startups were to both choose to recruit patients and collect data by themselves, it is yet another huge cost. In Taiwan, most patients are enrolled in the National Health Insurance program. Yet, hospital visits for startup administrated tests and/or general health examinations are not reimbursable under such a national program. Hence, startups often have to cover the entireties of such tests and/or examinations' fees for the participating patients. And most startups need to buy additional private insurance for such patients. Likewise, when hospitals cannot afford the test and decide to hire medical assistants to aid the project, startups have to pay such an assistant's salary. In Group B, a physician said that at a private health examination center, single patient data costs roughly 1 to 2 USD. Thus, a database with a total of roughly

10,000 patients' data would cost at a minimum of 10,000 USD. And such a cost does not include the other extraneous administrative and diagnosis expenses. Another startup involved physician said that such startup had paid roughly 830,000 USD in total to hospitals in a two-year patient recruiting project as reflected here: "*For most startups not even knowing how to pay their members for the next meal, the cost is mounting to astronomical figures*."

4.2.3 Other regulatory barriers

For startup company workers, accessing patient data is very difficult due to existing stringent regulations and a hospital's ethics committee. For example, C06 stated "*The IRB's regulation is too strict (for accessing patient data). It seemed like the regulation is only for protecting the IRB itself, instead of offering help or guarantee for startups*." At the onset, each organization with health databases has its own IRB. And each IRB has both different steps and regulations to follow. Thus, each time startups apply for a database, must be conformed to are different regulations and regulatory regimes. Especially for Taiwan's case, such conformities are common difficulties given Taiwan's dramatically different regulatory regime (civil law system-based) compared to those of other Western-based jurisdictional regimes that are often common law-based. Also, current IRBs often require every project applying for its approval to be a research or academic project. Thus, there is no concrete rule about how startups can apply, so the startups have to modify their project to somewhat reflect research. For startups, this is a task without any benefit or meaning that wastes considerable time and effort.

4.3 Desired infrastructure of accessing data for BMI startups

For the desired infrastructure, collected were BMI startups' feedback on existing infrastructures as shown in Table 3.

About regulations may be strict but well-informed and coherent: If BMI startups choose to both access and collect data through both hospital personnel's assistance and collaboration, most of them follow the IRB's at-hand rules. Yet, when BMI startups are applying for health data access without personal relationships with physicians, they in fact do not know which regulations to follow or where to start, nor can they find relevant regulations in order to proceed (e.g., A01, B02, B03, C04). Also we observed that participants expect the regulations to be both sound and coherent, so startups are still able to legally apply for data access if all regulations were to be followed.

About medical personnel: Group B's physician participants (B01, B02, B03) held similar attitudes to different extents. For example: B03 expected the regulations to be loosened. Yet, B03 said that data related to personal identity cannot be released without patient consent. B02 stated that, from a medical personnel's view, startups are profit-pursuing organizations. Thus, physicians have no right or reason to use patients' data to aid the physicians. And when patients visit hospitals for health examinations, treatments, etc., such patients did not know that their recorded data had additional uses. Thus, such data's release could be viewed as an ethical violation.

What's more, B01 anticipated that regulations will be eventually modified. At the same time, B01 believes that such relevant regulations can be

Table 3. Desired Data Accessing Infrastructure Perceived by BMI Startups

Aspects	Current infrastructure	Ideal infrastructure	Source
Regulation related to patient data access.	Strict and scattered.	Strict but connected; well-informed and transparent; coherent if regulations followed; the third party outside the IRB can authorize a "fastlane" for BMI startups.	A01, A04, B02, B03, and C04.
Data access opportunity.	Strict and scattered; centralized.	Possibility of exchange, selling, purchasing, accessing data in a decentralized, secure and safe approach – even if the regulation is rather strict.	C07
IRB and/or ethical review-related board.	Use both the same regulation and standard to review each product.	Different product types should have different reviewing processes. And the board can more efficiently review the application.	A01, B03, and C04.
BMI Startup Accelerator Program.	Provide a training course about BMI startups' data accessibility issues.	Besides training courses, the accelerator can mentor startups to go through the review process or provide a network to match startups' relevant needs.	A01 and C01.

made clearer so that startups can have a standard or guideline to follow. And B01 believes there will be solutions to both data-sharing and data-transferring issues as reflected here: "*The economy will die because of the restricted environment… when I am mentoring students in hospitals, I look at the new-coming doctors and expect they can understand they don't need to do every detailed work and can tell what can be achieved with data and technologies in the future*." Also, B01 mentioned that there were no personal worries about the regulations became stricter. Yet, B01 believed such regulations should be more well-connected, transparent, know what is profited, and what is allowed – for example, the United States' Health Insurance Portability and Accountability Act (HIPAA).

About tensions between the IRB approval process and startup development cycle: As a trusted third party, IRB has a vital responsibility to protect research subjects that include patients. To maintain high ethical standards, differing hospital IRBs also rigorously examine each other's data release decisions. Whereas, as suggested by some interviewees, months of the review process might be too long in time for startups, especially for enrolment, thereby slowing down a startup's technical innovation. For example: B02 stated that the current IRB might consider opening a "fast-track" for BMI startups to secure their respective technique and market advantage. Thus, it is fundamentally challenging to resolve such tensions between the IRB approval process and the startup development cycle. From a policy-making view, more research is needed in exploring possible solutions.

5. Discussion

5.1 Do I need raw data? Startups' perceptions of data needs

Overall, our results revealed differing types of discrepancies between both perceived and actual data needs. Namely as some Group A participants do not distinguish statistical information from raw data. For example, Participant P03 expressed that their startup would like to be able to access to the raw data in the NHIRD database. Yet, with further conversations, likewise found was that might be needed are some more concrete statistical facts (e.g., percentage of females who use software applications for monitoring their respective period and ovulation). Also arguable is that some participants who work in BMIs lacked the proper practices with a well-rounded training in mind in terms of both experiment design and the experience in dealing with research data. Namely as misconceptions might occur with such data requests. As in P03's case, the survey with such aforesaid female user surveys might in fact fulfill the IRB's exemption application in which there can be a relatively faster process compared to that of a regular application. With all these in mind, future research might assess the participants' data literacy and inquire more about data plus data privacy definitions, coupled with triangulation with respective data needs.

5.2 Suggestions for both government and startup accelerators

Also observed was that startups have strong demands for the government to make regulatory adjustments to ones that are both more transparent and well-informed: clearly informing startups what they both can and cannot do. For example: Most Group A participants expect a better channel for reaching potential users or self-reported patient data (e.g., patient-matching website). Namely as startups revealed that it takes two to three months for the IRB for approval. Such a time is a relatively long time in the developmental phase, namely for startups that apply for more than one hospital's health data. Also, the concerned authority can reconsider if the products can be managed by different levels, instead of requesting all the products to go through the same strict process. For startup accelerators, suggested are more focal points on how to assist BMI startups going through the relevant FDA or IRB applications. As aforesaid, startups might not be familiar with every applying regulation, or they may not know where their product falls under the relevant FDA classification at-hand. Thus, it would be very helpful if startup accelerators could provide mentorship on such complex regulations. And it is rather not easy for startups without medical personnel in their teams to reach physicians to serve as such startup projects' principal investigators. As startup accelerators might have a larger network of professionals from various relevant fields, perhaps can be introduced or recommend are startup development-interested physicians.

6. Conclusion

This study analyzed the current practices of healthcare startups accessing and collecting health data through interviews of 17 participants who worked in startups and/or as physicians through an in-depth interview approach. With such in-person interviews, identified were several obstacles that these startups encountered in balancing

development and ethics. Also summarized were such startups' health data accessibility expectations. And despite several details about unsatisfactory current practices, observed were that many of such obstacles can be concretely overcome altogether by the government, hospitals, and startup accelerators. Likewise, such findings are expected to provide deeper insights on building a mutually beneficial model for many stakeholders—thereby helping create a better environment for startups to both access and collect data. For future work, a larger systematic investigation can be proposed about BMI startup communities' various stakeholders that include respective personnel in the: information technology (IT) department, hospital legal affairs, BMI startups' legal affairs, and the public sector. Thus, researchers who are interested in this topic can be better rendered a more comprehensive outlook of the entire Taiwanese BMI ecosystem.

Ethics

The study was reviewed by the Research Ethic Office (equivalent to Institutional Review Board worldwide) at National Taiwan University in Taiwan and meets all the necessary criteria for minimal risk review (no. 201711HS017) in May 2018.

Acknowledgement

This work was financially supported by the Ministry of Science and Technology (MOST) in Taiwan, under MOST 110-2636-H-002-001- and MOST 110-2634-F-002-045, and the Center for Research in Econometric Theory and Applications (Grant no. 110L900204) from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan.

References

Bindley, K. (2019, November 22). Your health data isn't as safe as you think; silicon valley's rush into the health-care business is challenging the antiquated protections of Americans' medical histories. *Wall Street Journal (Online)*. Retrieved from https://search.proquest.com/docview/2316736717?accountid=14229

Boyacioglu, C., & Yıldız, O. (2021). The impact of data protection regulations on start-up enterprises. In O. Yıldız (Ed.), *Recent developments in individual and organizational adoption of ICTs* (pp. 120-133). doi: 10.4018/978-1-7998-3045-0.ch007

Carroll, A., & Frakt, A. (2017, December 26). The leap to single-payer: What Taiwan can teach how one nation transformed a health care system. *The New York Times*. Retrieved from https://www.nytimes.com/2017/12/26/upshot/the-leap-to-single-payer-what-taiwan-can-teach.html

CEOWORLD Magazine. (2021, April). *Revealed: Countries with the best health care systems, 2021*. Retrieved from https://ceoworld.biz/2021/04/27/revealed-countries-with-the-best-health-care-systems-2021/

Chang, C. H. (2016). Controversy over information privacy arising from the Taiwan National Health Insurance Database examining the Taiwan Taipei High Administrative Court Judgment No. 102-Su-

36 (Tsai v. NHIA). *Pace International Law Review, 28*(1), 29-116.

Chuang, Y., & Tsao, Y. (2012). Discussion on privacy protection technology of medical database (translated). *Journal of Information Communication & Technology Auditing, 25*, 30-43. doi: 10.30007/JICTA.201201.0003

Economist Intelligence Unit. (2017). *Global access to healthcare building sustainable health systems*. Retrieved from https://eiuperspectives.economist.com/sites/default/files/Globalaccesstohealthcare-3.pdf

Green, A. K., Reeder-Hayes, K. E., Corty, R. W., Basch, E., Milowsky, M. I., Dusetzina, S. B., ... Wood, W. A. (2015). The project data sphere initiative: Accelerating cancer research by sharing data. *The Oncologist, 20*(5), 464-e20. doi: 10.1634/theoncologist.2014-0431

Mack, N., Woodsong, C., MacQueen, K. M., Guest, G., & Namey, E. (2005). *Qualitative research methods: A data collectors field guide*. Retrieved from https://www.fhi360.org/sites/default/files/media/documents/Qualitative%20Research%20Methods%20-%20A%20Data%20Collector's%20Field%20Guide.pdf

Martin, N., Matt, C., Niebel, C., & Blind, K. (2019). How data protection regulation affects startup innovation. *Information Systems Frontiers, 21*(6), 1307-1324. doi: 10.1007/s10796-019-09974-2

Open Knowledge Network. (n.d.). *Global data index – by places*. Retrieved from https://index.okfn.org/place/

Patra, P., & Sleeboom-Faulkner, M. (2009). Bionetworking: Experimental stem cell therapy and patient recruitment in India. *Anthropology & Medicine, 16*(2), 147-163. doi: 10.1080/13648470902940689

Sudlow, R., Branson, J., Friede, T., Morgan, D., & Whately-Smith, C. (2016). EFSPI/PSI working group on data sharing: Accessing and working with pharmaceutical clinical trial patient level datasets – a primer for academic researchers. *BMC Medical Research Methodology, 16*(1), 23-36. doi: 10.1186/s12874-016-0171-x

Strong, D. M., Lee, Y. W., & Wang, R. Y. (1997). Data quality in context. *Communications of the ACM, 40*(5), 103-110. doi: 10.1145/253769.253804

(Received: 2020/10/28; Accepted: 2021/7/26)

圖書資訊學刊　第19卷第2期　（2021.12）　頁25-41
doi: 10.6182/jlis.202112_19(2).025

生物科技與醫藥新創產業社群中的病患資料近用性

Patient Data Accessibility for Biotech and Medicine Industry Start-ups in Taiwan

鄭　瑋[1]　王姿涵[2]　張維庭[3]

Wei Jeng[1], Zih-Han Wang[2], Wayland Chang[3]

摘　要

臺灣生物科技與醫藥產業環境中，病患資料的近用性（patient data accessibility）涉及倫理與法規上的考量，攸關資料釋出與取用上的供需平衡。本研究透過質化方法，對17位受訪者進行半結構式訪談（包含4位醫師、5位BMI新創公司人員，餘8位身兼上述兩種角色），以探索臺灣BMI新創社群在病患資料取得過程中，產生的可近用性需求與感知的困難。本研究發現，臺灣BMI新創產業環境中，病患資料取得成本相當高：除貨幣（金錢）成本之外，尚包括相關從業者因應生技法規之存在，而感知到必須付出的勞力與心力（perceived efforts）。研究結果亦顯示，臺灣的BMI從業者傾向期待政府能為醫患資料的取用建立更為完善、嚴明（而非放寬）的監管制度。

關鍵字：健康資料、資料近用與資料再用、資料基礎建設、生物科技與醫藥新創產業

[1,3]國立臺灣大學圖書資訊學系
Department of Library and Information Science, National Taiwan University, Taipei, Taiwan
[2] 美國華盛頓大學資訊學院
Information School, University of Washington, Seattle, USA
* 通訊作者Corresponding Author: 鄭瑋Wei Jeng, E-mail: wjeng@ntu.edu.tw

註：本中文摘要由作者提供。
以APA格式引用本文：Jeng, W., Wang, Z.-H., & Chang, W. (2021). Patient data accessibility for biotech and medicine industry start-ups in Taiwan. *Journal of Library and Information Studies, 19*(2), 25-41. doi: 10.6182/jlis.202112_19(2).025
以Chicago格式引用本文：Wei Jeng, Zih-Han Wang, and Wayland Chang. “Patient data accessibility for biotech and medicine industry start-ups in Taiwan.” *Journal of Library and Information Studies* 19, no. 2 (2021): 25-41. doi: 10.6182/jlis.202112_19(2).025