

Factors Related to Personal Health Data Sharing: Data Usefulness, Sensitivity and Anonymity

Samar HELOU^{a,1}, Victoria ABOU-KHALIL^b, Elie EL HELOU^c and Ken KIYONO^d

^aGlobal Center for Medical Engineering and Informatics, Osaka University, Japan

^bAcademic Center for Computing and Media Studies, Kyoto University, Japan

^cFaculty of Medicine, Saint Joseph University, Beirut, Lebanon

^dGraduate School of Engineering Science, Osaka University, Japan

Abstract. Using an online survey, we examined the relationships between the perceived usefulness, sensitivity, and anonymity of personal health data and people's willingness to share it with researchers. An analysis of 112 responses showed that people's willingness and perceptions are related to the type of the data, their trust in the data's anonymity, and their personal sociodemographic characteristics. In general, we found that people do not completely trust that their identities remain anonymous when sharing data anonymously with researchers. We also found that they are more willing to share personal health data with researchers if they perceive it as useful for public health research, not sensitive, and if they trust that their identity will remain anonymous after sharing it. We also found that people's age, gender, occupation, and region of residence may be related to their perceptions regarding the sharing of personal health data.

Keywords. personal health data, data sharing, willingness, trust, privacy, sensitivity

1. Introduction

Wide-scale sharing of Personal Health Data (PHD) is needed to enable data-driven healthcare. However, people are still reluctant to share their PHD for various reasons including privacy concerns, fear of discrimination, and the potential abuse of their data for commercial gain [1–3]. Peoples' willingness to share their PHD changes depending on the sensitivity of the data, the purpose of its use, the trustworthiness of the users, and the type of data and its safety [3–5]. Indeed, previous studies reported a higher willingness to share PHD with researchers rather than commercial entities [2–5] and indicated that people are willing to share their data when there is a clear public benefit. In addition, people are more reluctant to share sensitive health data, i.e., data that carries high personal risks if disclosed [1,5], and need to know that regulatory and technical safeguards exist when sharing their PHD [3].

Previous studies highlighted the relationship between the willingness of people to share their PHD and their perceived usefulness, sensitivity, and safety of the data. However, no previous study attempted to quantify these relationships and examine the interplay between these three factors. To address this knowledge gap, we aim to answer

¹ Corresponding Author, Samar Helou, E-mail: samar@bpe.es.osaka-u.ac.jp

the following question: *How do the perceived usefulness, sensitivity, and anonymity of personal health data relate to peoples' willingness to share it with researchers?*

2. Methods

Our aim was to assess the relationships between the perceived usefulness, sensitivity, and anonymity of PHD and people's willingness to share it with researchers. To do so, we first conducted a literature review aimed at identifying different types of health data that are usually contained in personal health records. Afterwards, through assessments and discussions between authors, we selected ten different types of personal health data with potentially varying degrees of sensitivity to include in the survey.

The survey was created in English using the SurveyMonkey platform and contained five demographic questions (age, gender, country of residence, occupation, and IT literacy) and 31 100-point continuous linear scale questions that examined the respondents' (i) willingness to share data; (ii) perception of the sensitivity of the data; (iii) perception of the usefulness of the data; (iv) perception of the anonymity of the data when sharing it anonymously with researchers. These perceptions were examined in regard to the ten different types of personal health data. The survey was tested and refined by two researchers and was finally administered as an anonymous online survey on social media platforms using a snowball sampling approach initiated by the paper's authors. The data was collected over the first week of December 2020. A copy of the survey can be provided upon request.

The collected data was analyzed using descriptive and inferential statistics and p-values less than 0.05 were considered statistically significant.

3. Results

We collected a total of 112 responses. The sample's demographics are shown in Table 1. The average age was $M = 30.32$, $SD = 10.37$, and the average self-reported IT literacy was $M = 66.95$, $SD = 17.95$, with "0 = I have no computer skills" and "100 = I can train and deploy machine learning models."

Table 1. The respondents' demographics

Characteristics		Frequency	Percentage
Gender	Female	64	57.1%
	Male	46	41%
	Other	2	1.8%
Region of residence	Middle East	68	60.7%
	Asia	15	13.4%
	Europe	16	14.3%
	North America	8	7.1%
	Oceania	3	2.7%
	South America	2	1.8%
Occupation	Healthcare	44	39.3%
	Research	17	15.2%
	IT	14	12.5%
	Other	37	33%

3.1. The willingness to share personal health data, and its perceived usefulness, sensitivity, and anonymity

On average, people reported slightly trusting that their identity remains anonymous if they share data anonymously with researchers (M = 65.03, SD = 26.07). Figure 1 shows how the willingness to share PHD, its perceived usefulness for public health research, and its perceived sensitivity differ for the 10 types of PHD.

The results show that height was the type of PHD that people were most willing to share with researchers (M = 92.24, SD = 16.99). Height was viewed as the least sensitive (M = 31.35, SD = 34.62), and as the least useful for public health research (M = 60.57, SD = 31.35). On the other hand, mental health records were the type of PHD that people were least willing to share (M = 67.3, SD = 30.04). Mental health records were viewed as the most sensitive (M = 77.95, SD = 28.28), but were perceived to have high usefulness to public health research (M = 82.9, SD = 22.86).

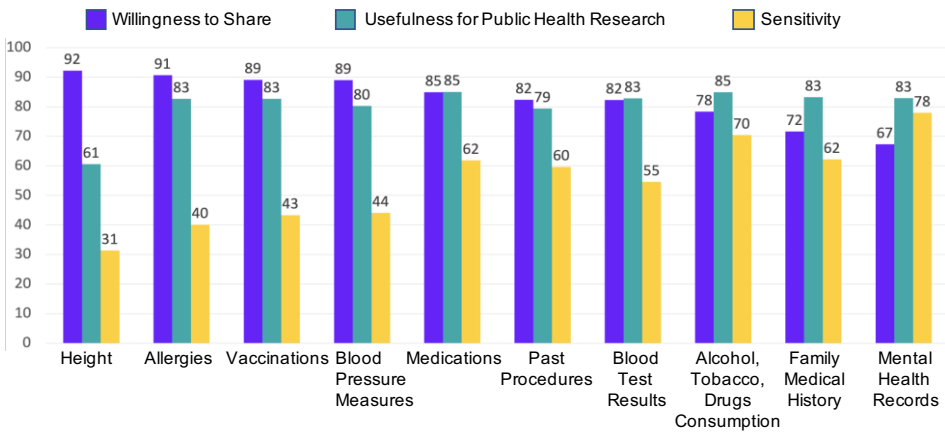


Figure 1. Willingness to share different personal health data and its perceived usefulness and sensitivity

Table 2 shows the Pearson correlations between the perceptions on sharing PHD. The results show that there are significant weak to moderate correlations between the willingness to share PHD with researchers and the data’s perceived usefulness for public health research (r = 0.253, p < 0.001), its perceived sensitivity (r = -0.334, p < 0.001), and its perceived level of anonymity once it is shared with researchers (r = 0.336, p < 0.001). There was also a weak, but statistically significant, correlation between the perceived anonymity of the data and its perceived usefulness for public health research.

Table 2. Relationship between willingness to share personal health data, and its perceived usefulness, sensitivity, and anonymity

	Willingness to share	Perceived usefulness	Perceived sensitivity	Perceived anonymity
Willingness to share	r = 1	r = 0.253* p < 0.001	r = -0.334* p < 0.001	r = 0.336* p < 0.001
Perceived usefulness	r = 0.253* p < 0.001	r = 1	r = 0.056 p = 0.062	r = 0.153* p < 0.001
Perceived sensitivity	r = -0.334* p < 0.001	r = 0.056 p = 0.062	r = 1	r = -0.028 p = 0.344
Perceived anonymity	r = 0.336* p < 0.001	r = 0.153* p < 0.001	r = -0.028 p = 0.344	r = 1

3.2. Sociodemographic factors and perceptions on sharing personal health data

3.2.1. Age and IT literacy

We found weak, but significant, correlations between age and perceptions on sharing PHD. Moreover, self-reported IT literacy was weakly correlated with the perceived anonymity of shared PHD. No correlation was found between age and IT literacy. The Pearson correlation coefficients are shown in Table 3.

Table 3. Relationship between age and self-reported IT literacy, and perceptions on sharing PHD

	Willingness to share	Perceived usefulness	Perceived sensitivity	Perceived anonymity
Age	r = -0.188* p < 0.001	r = -0.137* p < 0.001	r = 0.161* p < 0.001	r = -0.105* p < 0.001
IT literacy	r = 0.027 p = 0.375	r = 0.051 p = 0.09	r = 0.050 p = 0.094	r = 0.109* p < 0.001

3.2.2. Gender, occupation, and region of residence

A three-way MANOVA showed significant main effects and interaction effects of demographic factors (occupation, region of residence, and gender) on peoples' perceptions regarding the sharing of their PHD. The effects are shown in Table 4. The following subsections highlight the group differences that were found through post hoc comparisons using the Tukey HSD test.

Table 4. Wilks Lambda test results showing the main and interaction effects of the three demographic factors

Effect	Value	F	Hypothesis df	Error df	Sig.
Occupation	0.883	8.623	16.000	3321.477	<0.001
Region	0.747	16.537	20.000	3606.121	<0.001
Gender	0.937	8.996	8.000	2174.000	<0.001
Occupation*Region	0.844	6.758	28.000	3920.656	<0.001
Occupation*Gender	0.940	5.680	12.000	2876.223	<0.001
Region*Gender	0.905	6.929	16.000	3321.477	<0.001
Occupation*Region*Gender	0.984	2.197	8.000	2174.000	0.025

3.2.2.1. Occupation

Healthcare workers (M = 88.11, SD = 19.77) were significantly more willing to share their PHD than IT workers (M = 72.16, SD = 28.89), researchers (M = 78.26, SD = 27.58), and workers in other fields (M = 82.55, SD = 27.84). Interestingly, researchers (M = 57.76, SD = 28.44) and people who do not work in healthcare, IT, or research fields (M = 58.03, SD = 28.44) were significantly less trusting than healthcare workers (M = 71.98, SD = 22.16) and IT workers (M = 70.5, SD = 19.15) that their identities remain anonymous when sharing their PHD with researchers.

3.2.2.2. Region of residence

People in Europe (M = 87.2, SD = 21.25) and the Middle East (M = 87.34, SD = 22.67) were significantly more willing to share their data than people in North America (M = 61.15, SD = 25.43), and Asia (M = 68.16, SD = 32.2) and perceived the sensitivity of their PHD as significantly lower, and its usefulness significantly higher. However, people in Europe (M = 76.06, SD = 20.33) and North America (M = 74.13, SD = 24.83) trusted that their identity remains anonymous significantly more than people in the Middle East (M = 65.01, SD = 24.75) and Asia (M = 58.27, SD = 27.57).

3.2.2.3. Gender

The female and male groups were similar in terms of willingness to share, perceived usefulness, and perceived sensitivity of PHD. However, the female group had significantly more trust that the data remained anonymous ($M = 67.22$, $SD = 26.52$) than the male group ($M = 62.67$, $SD = 24.92$) and the group identifying as other ($M = 49.00$, $SD = 21.55$).

4. Discussion

Our results confirmed that not all PHD are perceived equally, and that people are more willing to share their PHD with researchers if they perceive it as not sensitive, useful for public health research, and trust that their identities remain anonymous after sharing it. However, people do not completely trust that their identity remains anonymous. These anonymity concerns must be addressed if we aim to increase people's willingness to share their PHD. Interestingly, we found no relationship between the trust that PHD remains anonymous and the perceived sensitivity of the data, but both were positively correlated with the willingness to share PHD. This means that people view the data's level of sensitivity and its anonymity as independent factors. Therefore, in cases where people are being asked to share sensitive PHD, promises of anonymity would not be enough. In these cases, showing how this data can be useful for public health research may increase people's willingness to share it.

We also found initial evidence that sociodemographic characteristics are related to perceptions of, and willingness to share, PHD. However, our sample is small and limited. Further research is needed to examine the perceptions and concerns of different people.

5. Conclusions

People are more willing to share their personal health data with researchers if they perceive it as useful for public health research, not sensitive, and trust that they cannot be identified after sharing it. Moreover, people may have different perceptions and levels of concerns regarding the sharing of personal health data depending on their socio-demographic characteristics, with age, gender, and occupation being potential factors. This study's results confirmed previous findings and highlighted aspects to be considered to increase people's willingness to share their personal health data with researchers.

References

- [1] Weitzman ER, Kelemen S, Kaci L, et al. Willingness to share personal health record data for care improvement and public health: a survey of experienced personal health record users. *BMC medical informatics and decision making*. 2012;12:1–10.
- [2] Pickard KT, Swan M. Big Desire to Share Big Health Data: A Shift in Consumer Attitudes toward Personal Health Information. *AAAI Spring Symposia*. 2014.
- [3] Mori I. The one-way mirror: public attitudes to commercial access to health data. Wellcome Trust; 2017.
- [4] Karampela M, Ouhbi S, Isomursu M. Connected Health User Willingness to Share Personal Health Data: Questionnaire Study. *Journal of medical Internet research*. 2019;21:e14537.
- [5] Soni H, Grando A, Murcko A, et al. State of the art and a mixed-method personalized approach to assess patient perceptions on medical record sharing and sensitivity. *Journal of Biomedical Informatics*. 2020;101:103338.