HIR
Healthcare Informatics Research

Privacy Protection and Data Utilization

Soo-Yong Shin[1,2,3]
[1]Department of Digital Health, SAIHST, Sungkyunkwan University, Seoul, Korea
[2]Department of Intelligent Precision Healthcare Convergence, Sungkyunkwan University, Suwon, Korea
[3]Big Data Research Center, Samsung Medical Center, Seoul, Korea

To balance privacy protection and data usage, the Personal Information Protection Act (PIPA) in Korea was amended last year [1]. Based on the amended PIPA, the personal data pseudonymization guideline by the Personal Information Protection Commission and the medical data utilization guideline by the Ministry of Health and Welfare and the Personal Information Protection Commission were also published [2]. However, there are still many debates on how to use medical data due to the remaining unclear regulations.

For example, the definitions of medical data and health-related data are unclear in the regulations. Although more than 19 regulations, including the Medical Service Act [3] and the Bioethics and Safety Act [4], mention the medical data, there are no clear definitions. Article 3(6) of the Framework Action on Health and Medical Service defines "information on health and medical services" as knowledge or all kinds of data expressed in the form of codes, figures, letters, voice, sound, images, and so forth, which are related to health and medical services [5]. However, this definition is too broad and repetitive.

In particular, PIPA defines sensitive information, which should be more carefully protected compared to other personal information. Sensitive information includes "health" without any formal definition of what health is. As technologies develop, the domain of health-related data is sharply expanding. Therefore, this gives rise to controversial points.

The other issue is the relationships among regulations. Most researchers are still confused about the priority of regulations when using health data. The purpose described in Article 1 of Medical Service Act is "to provide for the matters necessary for the provision of medical services to people to ensure that people can enjoy the benefits of high-quality medical treatment" [3]. This is the primary purpose of utilizing clinical data. Scientific research is the secondary purpose of clinical data. Therefore, the Bioethics and Safety Act, not the Medical Service Act applied to scientific research.

If we clearly define the purpose of using clinical data (clinical practice vs. clinical research), we can easily identify which act should be applied. For example, the Medical Service Act applied to data sharing for clinical practice; on the other hand, the Bioethics and Safety Act applied to data sharing for research or product development. If medical records are pseudonymized, the pseudonymized clinical data will be affected by the PIPA, not the Medical Service Act [2], because pseudonymized data cannot be used for clinical (primary) purposes.

However, because pseudonymization in PIPA should be regarded as one of the methods of anonymization in the Bioethics and Safety Act based on the authoritative interpretation in the guideline [2], all studies using pseudonymized clinical data should be approved by an Institutional Review Board (IRB). There are claims that pseudonymized data for scientific research could be used freely based on the amended PIPA; however, this is incorrect. All research using human subjects should be approved by an IRB due to the Bioethics and Safety Act. IRB approval could be time-consuming. However, we must keep in mind that IRB approval is a minimal safeguard to protect both patients' rights and

researchers' rights.

In addition, many researchers have tried to implement ethical artificial intelligence (AI), and the Korean government has published ethics guideline for AI. Prior to implementing ethical AI and utilizing AI in ethical ways, studies themselves should be ethical. Many technical points have been changed based on the amended PIPA and the guidelines. However, nothing has been changed in the fundamental process of human subject research.

Many researchers claim that the newly published medical data utilization guideline is too strict for both research and business purposes. However, it should be noted that we have a first step to move forward. There is no way to satisfy everyone. Based on the guidelines, we have to implement institutional regulations and accumulate real experiences to amend the related regulations. At the same time, we could try to find alternatives to pseudonymization because pseudonymization or anonymization methods basically distort data to hide identities. Distorted data can distort research results.

Privacy-preserving data mining techniques or technical solution for consent could be alternative approaches. Privacy-preserving data mining techniques analyze data while satisfying privacy protection requirements; examples include homomorphic encryption or federated learning. Alternatively, using synthetic data, which are fake data generated from real data, could be another solution. Technical solutions regarding consent, such as dynamic consent, could provide easier ways to collect consent from research participants.

It is very hard to satisfy both privacy protection and data utilization needs. However, we, as researchers, should try to find suitable technical solutions to satisfy the current regulation as well as to amend the regulation based on the real world experience. Also, we should keep in mind that pseudonymized data is a kind of personal information as stated in Article 2(1) of PIPA. Therefore, we must protect pseudonymized data in the same way as we must protect personal information.

ORCID

Soo-Yong Shin (https://orcid.org/0000-0002-2410-6120)

References

1. Korea Legislation Research Institute. Personal Information Protection Act [Internet]. Seoul, Korea: Korea Legislation Research Institute; 2020 [cited at 2021 Jan 13]. Available from: https://elaw.klri.re.kr/eng_service/lawView.do?hseq=53044&lang=ENG.

2. Ministry of Health and Welfare. Medical data utilization guideline [Internet]. Seoul, Korea: Korea Legislation Research Institute; 2020 [cited at 2021 Jan 13]. Available from: http://www.mohw.go.kr/react/al/sal0101vw.jsp?PAR_MENU_ID=04&MENU_ID=040101&page=1&CONT_SEQ=360055.

3. Korea Legislation Research Institute. Medical Service Act [Internet]. Seoul, Korea: Korea Legislation Research Institute; 2020 [cited at 2021 Jan 13]. Available from: https://elaw.klri.re.kr/eng_service/lawView.do?hseq=53532&lang=ENG.

4. Korea Legislation Research Institute. Bioethics and Safety Act [Internet]. Seoul, Korea: Korea Legislation Research Institute; 2020 [cited at 2021 Jan 13]. Available from: https://elaw.klri.re.kr/eng_service/lawView.do?hseq=46341&lang=ENG.

5. Korea Legislation Research Institute. Framework Act on Health and Medical Services [Internet]. Seoul, Korea: Korea Legislation Research Institute; 2018 [cited at 2021 Jan 13]. Available from: https://elaw.klri.re.kr/eng_service/lawView.do?hseq=49490&lang=ENG.