

Review

Not peer-reviewed version

Clarifying Guidelines for the Use of Medical Data for Research in South Korea: Reconciling the Bioethics and Safety Act and Personal Information Protection Act

[Jae Sun Kim](#) *

Posted Date: 5 March 2024

doi: 10.20944/preprints202403.0232.v1

Keywords: medical data; human-subject research; scientific research; consent; PIPA; BSA



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Review

Clarifying Guidelines for the Use of Medical Data for Research in South Korea: Reconciling the Bioethics and Safety Act and Personal Information Protection Act

Jae Sun Kim ^{1,*}

¹ College of Law, Dongguk University-Seoul, 04620 Seoul, Republic of Korea; jaesk910@gmail.com; jaeskim@dongguk.edu.

* Correspondence: jaesk910@gmail.com; Tel.: 82-2260-85651

Abstract: The 2020 revision of South Korea's Personal Information Protection Act (PIPA) updated procedures for the use of medical data for research. The PIPA provision allowing "pseudonymized information" to be used without the consent of the data subject for scientific research introduced ambiguity regarding the applicability of the PIPA versus the Bioethics and Safety Act (BSA), which allows "anonymized information" to be used for research only after data-subject consent and IRB approval for human-subject research are obtained. The main points requiring clarification include the distinction between "human-subject" and "scientific" research, the definition of anonymization and pseudonymization, the requirements for data-subject consent, and the procedures for institutional review boards (IRBs) and data review boards (DRBs). After reviewing the legal concepts and processes, this article recommends clarification of the guidelines for the use of medical data in research by distinguishing the concept and, in the long run, categorizing medical data based on the degree of risk.

Keywords: medical data; human-subject research; scientific research; consent; PIPA; BSA

1. Introduction

In 2020, South Korea revised the Personal Information Protection Act (PIPA), introducing the concept of "pseudonymized information" that could be processed without the consent of data subjects for scientific research [1]. However, the applicability of the new PIPA procedures and the pre-existing, and different, Bioethics and Safety Act (BSA) procedures for conducting human-subject research were never deconflicted, resulting in considerable confusion in practice [2]. Discussions emerged regarding the different purposes of human-subject research and scientific research, the definition of anonymization and pseudonymization in terms of information processing, consent in terms of information utilization, and the procedures for institutional review boards (IRBs) and data review boards (DRBs).

Requirements for human-subject research under the BSA—namely, IRB approval and specific consent from data subjects—are applied quite strictly in South Korea. However, with the rapid development of information processing technologies, and especially after the outbreak of COVID-19 in 2020, the demand to make medical data available for research purposes has increased not only from researchers from medical institutions or governments but also from diverse industries [3]. In particular, when the PIPA revision introduced procedures for utilizing pseudonymized information without consent, it sparked heightened interest in the use of medical data for research—a use that had previously only been authorized through consent and IRB procedures.

This article examines the current statutes regarding the utilization of medical data for research, including the BSA and PIPA, and proposes legal revisions to clarify the guidelines for the use of medical data for human-subject or scientific research.

2. Results and Discussion

2.1. Increased Demand for the Use of Medical Data for Research

Medical institutions have long collected and utilized medical data for the purpose of medical practice such as the diagnosis and treatment of disease. The BSA regulates the conduct of human-subject research as an extension of medical practice and stipulates the requirement for consent of the data subjects and IRB approval for the sake of bioethics and safety.

The value of health data is rapidly evolving, however, due to advances in AI language-processing technologies such as application programming interfaces (API), which are healthcare natural language tools that can analyze electronic medical records (EMR) and electronic health records (EHR); next-generation sequencing (NGS) technologies that can analyze genomic DNA sequences; and AI technologies that analyze medical images such as chest X-rays to identify the probability of cancer or lung disease [4,5]. For instance, in South Korea, start-up company Lunit introduced Lunit Insight CRX, which utilizes AI technology to diagnose thoracic diseases such as lung cancer and lung nodules, and Lunit Insight MMG, which analyzes images from mammograms to detect the presence of breast cancer with an accuracy rate of 96% [6].

2.2. Laws on the Use of Medical Data for Research

In South Korea, the Medical Service Act[7] and BSA mainly apply to medical institutions and researchers, and human-subject research can only be conducted after receiving IRB approval and consent from the subject. However, the 2020 PIPA revision permits the use of pseudonymized personal information for scientific research purposes without data-subject consent, which has resulted in legal confusion regarding the distinctions between “human-subject” (covered by the BSA) and “scientific” research (which falls under PIPA rules), the definitions of anonymization and pseudonymization, and the deliberation and operation of IRBs and DRBs [8].

Table 1. BSA v. PIPA.

		BSA	PIPA
Purpose		Governs research on human beings, human materials, etc., handling embryos, genes, etc.	Processing and protection of personal information
Values		Human dignity and values, bioethics and safety	Freedom, rights, dignity, and value of individuals
Subject		Human-subject research	Scientific research
Measures		Anonymization	Pseudonymization
Use	Identifiable	Consent + IRB → human subject research (Arts. 15 & 16) * regulated without distinction	(1) Use only within purpose; (2) Out-of-purpose use allowed if (i) consented, (2) under other legal grounds, (3) imminent danger (Art. 18; Art. 15 of the Enforcement Decree)
	Anonymous/ Pseudonymous		Pseudonymization + DRB (Guidelines) → scientific research (Art. 28-2)
Provision	Identifiable (provided as is)	Consent + IRB → human subjects research (Art. 18)	Consent (Art. 17)
	Provided after Anonymization	Consent + IRB + anonymization → human-subject research (Art. 18)	(Consent not needed) + pseudonymization + DRB → scientific research (Art. 28-2)

	/Pseudonymization		
Use/ Provision	Guidelines Use/Provision after Pseudonymization	IRB (research plan review) + exemptions if applicable [IRB exemption (i.e. when using existing information) or consent exemption (i.e. impracticality)] → pseudonymization + DRB → scientific research (Art. 28-2)	

* Guideline: Guidelines for Utilizing Health and Medical Data.

2.2.1. Human-Subject Research under the BSA, Scientific Research under the PIPA

The BSA mainly pertains to bioethics and safety; its purpose is “to ensure bioethics and safety by preventing the violation of human dignity and values or the infliction of harm on a human body in the course of research on human beings, human materials, etc. or of handling embryos, genes, etc.” The purpose of the PIPA is “to realize the freedom, rights, dignity, and value of individuals” by “prescribing requirements for the processing and protection of personal information.” Under Article 16 of the BSA, human-subject research requires consent from the human subjects and approval from the competent IRB. However, according to Article 13 of the BSA’s Enforcement Rule, the IRB review requirement is waived for “research that does not involve invasive activities but only simple measurement or observation, and research that does not identify the research subjects and does not collect or record sensitive information, regardless of whether the research is conducted face-to-face, but uses existing materials or documents”; additionally, when “consent is impracticable, poses a serious risk to the validity of the research, or the risk to the research subjects is extremely low,” consent is exempted according to Article 16 paragraph 3 of the Act.

When scientific research is conducted in accordance with the PIPA, personal information may be 1) utilized within the scope of the research purpose after consent to the use of personal information is obtained in accordance with Article 17 of the PIPA, 2) used for scientific research if the information is pseudonymized in accordance with Article 28-2 of the Act and the Guidelines and the appropriateness of pseudonymization is confirmed by the DRB, and 3) provided to third parties conducting scientific research.

Table 2. Human-subject Research Exempt under BSA.

Exempt from IRB (Art. 15; Art. 13 of the Enforcement Rule)	Exempt from Consent (Art. 16 para. 3)
<ul style="list-style-type: none"> - Research that does not involve invasive activities, research involving simple measurement or observation - Research that does not identify the research subjects and does not collect or record sensitive information - Research that uses existing materials or documents 	<ul style="list-style-type: none"> - When consent is impracticable - Poses a serious risk to the validity of the research - When risk to the research subjects is extremely low

Both the BSA and PIPA distinguish between human-subject research and scientific research; however, in practice, this distinction is opaque. Generally, human-subject research dealing with genes, embryos, etc., is prioritized over general personal information, and research utilizing data from medical institutions is subject to human-subject research procedures if it is judged that protection is warranted based on bioethics and values. However, if information is intended to be used for scientific research through pseudonymization, it may fall under the IRB exemption (Article 15 of the Act, Article 13 of the Enforcement Rule) or consent exemption (Article 16 paragraph 3 of the Act), allowing for simplified procedures. Therefore, distinguishing between “human-subject research” and “scientific research” is necessary.

Procedures for human-subject research using information that has already been anonymized or pseudonymized are not clearly defined; both the BSA and the PIPA stipulate that "a person who intends to conduct a human-subject research project shall prepare a research plan and submit it for review by the competent IRB before commencing such human-subject research," and that "a personal information controller may process pseudonymized information for scientific research purposes without the consent of data subjects"; this criteria applies based on the research purpose, not whether the existing information is pseudonymized or anonymized. Therefore, approval should be based on whether the purpose of the research is human-subject research or scientific research.

For human-subject research using already anonymized information, IRB review may be waived because the research is likely to be "non-invasive or utilize existing materials or documents." Consent may also be waived because it is likely to be "impractical to obtain consent, or the risk for the research subjects is extremely low." However, Article 15(1) of the BSA still requires IRB approval of the research plan before conducting human-subject research, and the IRB review may consider whether the research should be exempt from consent requirements or further IRB review.

2.2.2. Transfer of Medical Information to Third Parties under the BSA and the PIPA

According to Article 18 of the BSA, a human-subject researcher may provide data subjects' personal information to a third party after obtaining the consent of the data subjects under Article 16 of the Act and receiving IRB approval under Article 18. In accordance with Article 18 paragraph 2 of the Act, identifying information can be provided to a third party if the research subject has consented to such provision, but if the subject has not consented, then the information must be anonymized. According to Article 17 of the PIPA, the personal information processor may provide personal information to a third party if consent is obtained from the data subject and the personal information is provided within the scope of purpose for which it was collected. Further, according to Article 28-2 of the PIPA, personal information may be pseudonymized for scientific research without the consent of the data subjects and provided to a third party after DRB approval of pseudonymization in accordance with the Guidelines.

The BSA stipulates that human-subject research requires data-subject consent for the provision of information to a third party, IRB review, and anonymization, but PIPA rules for scientific research allow for the provision of pseudonymized information to a third party (with DRB approval of pseudonymization) without data-subject consent. The current Guidelines stipulate that human-subject research using pseudonymized information requires IRB approval of the research plan in accordance with the BSA and explain that exemptions are available if the research falls under "using existing materials or documents" under Article 13 of the Enforcement Rule [9]. However, the Guidelines do not clearly explain the rules for scientific research that falls under the PIPA.

2.3. Main Legal Issues with Using Medical Data for Research

2.3.1. Legal Distinction between Human-Subject Research and Scientific Research

As discussed above, human-subject research using anonymized information under the BSA and scientific research using pseudonymized information under the PIPA have different procedural and methodological requirements. Confusion arises because the scope of what constitutes human-subject research versus scientific research is not clearly distinguished.

The BSA defines human-subject research as "a research project specified by the Ordinance of the Minister of Health and Welfare, such as a research project physically involving a human being as a subject or conducted through communication, physical contact, or other means of interaction with a human being, or a research project conducted using information by which individuals could be identified." The Enforcement Rule of Bioethics and Safety uses a very similar definition, encompassing research that obtains data by directly manipulating the research subject or manipulating the research subject's environment (intervention research), research that obtains data by observing the behavior of the research subject or conducting face-to-face surveys (interaction research), and research that involves the use of information by which individuals can be directly or

indirectly identified (personal information use research). The PIPA defines scientific research as "research that applies scientific methods such as technological development and demonstration, fundamental research, applied research, and privately funded research," and includes "not only natural scientific research, but also historical research that applies the scientific method, research conducted in the public interest in the field of public health, as well as research for industrial purposes, such as research and development and improvement of new technologies, products, and services." Therefore, the classification of research using medical institutions' health data as human-subject research or scientific research is based on the research plan in terms of "whether there is physical involvement of human beings, communication and other means of interaction, or whether personally identifiable information is utilized" (in the case of human-subject research) or "when the scientific method is applied to technology development and demonstration, fundamental research, applied research, and privately funded research, etc."

Table 3. Human-Subject Research Definition.

BSA	Research project specified by Ordinance of the Minister of Health and Welfare, such as a research project physically involving a human being as a subject or conducted through communication, physical contact, or other means of interaction, and a research project conducted by using information with which individuals can be identified
Enforcement Rule of BSA (Art. 2 subpara.1)	<ul style="list-style-type: none"> - Research physically involving a human being (intervention research) - Research project involving communication, physical contact, or other means of interaction (interaction research) - Research using information with which individuals can be directly or indirectly identified (personal information use research)

Table 4. Scientific Research Definition.

PIPA (Art. 2 subpara. 8)	Research that applies scientific methods, such as technological development and demonstration, fundamental research, applied research, and privately funded research
Guidelines for Utilizing Health and Medical Data	Scientific research includes historical research that applies the scientific method, research conducted in the public interest in the field of public health, as well as research for industrial purposes, such as research and development and improvement of new technologies, products, and services

2.3.2. Interpretation of Pseudonymized Data for Scientific Research

For scientific research that utilizes pseudonymized medical data, if there is no "physical intervention, communication, or other interaction with humans," the relevant judgment is whether the pseudonymized information falls under the definition of "using information that could identify an individual" found in Article 2, paragraph 1 of the BSA's Enforcement Rule. Article 2 subparagraph 17 of the BSA defines "personally identifiable information" as "information by which an individual can be identified, such as the name and resident registration number of a human research subject or the donor of an embryo, oocyte, spermatozoon, or human material (hereinafter referred to as 'human research subject or donor')," and subparagraph 18 defines personal information as "information about an individual, such as personally identifiable information, genetic information, or information about health." Under Article 2 subparagraph 1(c) of the PIPA, pseudonymized information is defined as "information that is incapable of uniquely identifying an individual without the use or combination of information for restoration to its original state," so even if the definition in Article 2 paragraph 1 of

the BSA's Enforcement Rule is applied, such research is unlikely to fall under the definition of "research using personally identifiable information."

Table 5. Legal Definition of Information.

Personally Identifiable Information (Art. 2 subpara. 17 of the BSA)	Information by which an individual could be identified, including the name and resident registration number of a human subject of research or the donor of an embryo, oocyte, spermatozoon, or human material
Personal information (Art. 2 subpara. 18 of the BSA)	Information about an individual, including personally identifiable information, genetic information, or information about health
Pseudonymized information (Art. 2 subpara. 1(c) of the PIPA)	Information that is incapable of uniquely identifying an individual without the use or combination of information for restoration to its original state

However, the "Ethical Guidelines for Researchers by the Korea National Institute for Bioethics Policy"[10] defines "research using information that could identify individuals" as "research using information that could directly or indirectly identify research subjects (personal information used for research)." Pseudonymized information, by definition, does not fall under the scope of directly identifiable information because it is "processed so that a specific individual cannot be recognized without additional information," but it could be interpreted as "indirectly identifiable information," so determining the applicability of these regulations is complex [11].

The rule regarding what constitutes human-subject research is still open to interpretation in that pseudonymized information can, in principle, be identified when combined with other information and, thus, could be considered indirectly identifiable information. Another issue is balancing the interests of information utilization procedures. Under the BSA, anonymous information refers to "information that can no longer identify an individual in conjunction with other information with reasonable consideration of time, cost, technology, etc.," while pseudonymous information is defined on the premise that a specific person could be recognized by "using or combining additional information." Anonymized information is considered more secure and more strictly protective in terms of personal identification than pseudonymized information. However, if the information is provided for human-subject research purposes after anonymization, consent is exempted, and if the information is provided for scientific research purposes after pseudonymization, IRB review is required but consent is exempted. Therefore, a review is required to determine whether the balance of legal interests has been met [12,13].

Table 6. Legal Definition of Anonymized and Pseudonymized Information.

Anonymized information under the BSA	<u>Information that could no longer identify an individual in conjunction with other information with reasonable consideration of time, cost, technology, etc.</u>
Anonymization under the BSA	<u>The permanent deletion of personally identifiable information or full or partial substitution of personally identifiable information with an identification code given by a participating institution</u>
Pseudonymized information under the PIPA	<u>Information that has gone through a pseudonymization procedure and is incapable of uniquely identifying an individual without the use or combination of information for restoration to its original state</u>
Pseudonymization under the PIPA	A procedure to process personal information so that the <u>information cannot uniquely identify an individual without additional information</u> , by erasing in part, or replacing in whole or in part, such information

3. Conclusion: Key Considerations for Improving Legislation

3.1. *The Need for Distinction between Human-Subject and Scientific Research*

The scope of human-subject research requires a more precise definition to distinguish it from scientific research. Because the current definition of human-subject research includes information that directly or indirectly identifies an individual, a literal interpretation would include pseudonymized information, which could be re-identified by combining it with other information (even though there are penalties for re-identification). Therefore, the definition of human-subject research in the BSA's Enforcement Rule should be clarified to limit human-subject research to research that uses information that can directly identify individuals or exclude research that uses indirectly identifiable information.

Because the utilization of pseudonymized information in scientific research is permitted under the PIPA, the need for data-subject consent in the case of health data should be reconsidered. While human-subject research requires the data subject's consent even for the provision of anonymous information, scientific research does not require consent for pseudonymized information. While neither the current BSA nor PIPA define the information collection path, most health data collected in medical institutions is collected for the purpose of medical treatment rather than research, so consent is a key component of the data subject's information rights. The question of whether information collected for the purpose of medical practice (treatment) can be pseudonymized and provided to third parties for scientific research purposes without the consent of the data subject (as with general personal information) may need to be reconsidered in terms of protecting the data subject's rights. Regardless of whether legislation requires the same strict consent procedures as for human-subject research through anonymization, consideration should be given to ensuring notice and opt-out rights for data subjects even when health data are pseudonymized for scientific research [14,15].

For retrospective studies, in particular, while the information is used after anonymization and consent for the purposes of the initial studies, when the anonymized information is provided to a third party, the data is ultimately exempt from additional consent requirements because "it is impracticable to obtain consent." From the data subject's perspective, when anonymized information is provided for a specific study and is subsequently provided to a third party, additional consent cannot be given. Therefore, in such cases, even if it is impractical to legislatively require direct consent when providing anonymized information to a third party, consideration can at least be given to adding a public notification procedure to replace the notification to the data subject.

3.2. *The Need to Establish a System for Risk-Based Categorization*

Conflicting terminology in the legislative history of the BSA and PIPA leads to confusion in practice. In particular, while there is not much technical difference in the identifiability of anonymization and pseudonymization, there is a significant difference in the consent and utilization procedures between anonymization under the BSA and pseudonymization under the PIPA, and, thus, difficulties arise in practice when the Medical Service Act is applied. Regulations should be improved by categorizing health data according to the degree of risk, such as identifiability, to simplify the utilization process for low-risk information and clarify the process for high-risk information. In the same vein, the concepts of anonymization and pseudonymization within related processes need to be more clearly defined based on the degree of risk.

Author Contributions: conceptualization, methodology, validation, formal analysis, investigation, resources, writing, funding acquisition, J.S.K.

Funding: This work was supported by the 2022Dongguk University Research Fund.

Institutional Review Board Statement: Not Applicable.

Informed Consent Statement: Not Applicable.

Data Availability Statement: Data sharing not applicable.

Acknowledgments: The author thanks Dongguk University for supporting this research.

Conflicts of Interest: The author declares no conflict of interest.

References

1. PIPA, Article 28-2
2. In 2020, 3 major acts including the "Personal Information Protection Act, Credit Information Use And Protection Act, Act On Promotion Of Information And Communications Network Utilization And Information Protection" were revised to promote the use of data for research using "Pseudonymized Information".
3. Evans, M. Hospitals give tech giants access to detailed medical records. *Wall St. J.*, Jan. 20, 2020.
4. Sharon T.; Zandbergen, D. From data fetishism to quantifying selves: Self-tracking practices and the other values of data. *New Media Soc* 2017, 19, pp. 1695-1709.
5. Terry, N.P. Regulatory disruption and arbitrage in health-care data protection. *Yale J. Health Pol'y Law & Ethics* 2017, 17, pp. 143–208.
6. See news release, Lunit presents 11 research abstracts on AI image analysis solutions at the Radiological Society of North America. *Money Today*, Nov. 2, 2022.
7. Act No. 19818.
8. Kim, M.W.; Kim, I.H. A Study on Legislation for Protection and Use of Sensitive Information in the Intelligent Information Society, *Seoul Law Review*. 2021, 29-2, pp. 106-111.
9. Guidelines for Utilizing Health and Medical Data.
10. Ethical Guidelines for Researchers at the National Institute of Bioethics.
11. Ko, M.S. A Study on the Personal Information Protection in the Human Subject Research Project, *Legal Legislation* 2023, 701, pp. 20-23.
12. Ohm, P. Broken promises of privacy: Responding to the surprising failure of anonymization. *UCLA L.R.* 2010, 57, pp. 1701-1778.
13. Singer, N. Flo settles F.T.C. charges of misleading users on privacy. *New York Times*, January 13, 2021.
14. Stramm, J. Responding to the digital health revolution. *Rich. J.L. & Tech.* 2021, 128, pp. 86-160.
15. Huckvale, K. Assessment of the data sharing and privacy practices of smartphone apps for depression and smoking cessation, *JAMA Netw Open*. 2019, 2, e192542.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.