HIR
Healthcare Informatics Research

Current Status and Key Issues of Data Management in Tertiary Hospitals: A Case Study of Seoul National University Hospital

Jinwook Choi[1], Hyeryun Park[2], Eui Kyu Chie[3,4], Sae Won Choi[3], Ho-Young Lee[5,6], Sooyoung Yoo[5], Byoung Jae Kim[7], Borim Ryu[8]

[1]Department of Biomedical Engineering, Seoul National University College of Medicine, Seoul, Korea
[2]Interdisciplinary Program for Bioengineering, Graduate School, Seoul National University, Seoul, Korea
[3]Office of Hospital Information, Seoul National University Hospital, Seoul, Korea
[4]Department of Radiation Oncology, Seoul National University College of Medicine, Seoul, Korea
[5]Office of eHealth Research and Businesses, Seoul National University Bundang Hospital, Seongnam, Korea
[6]Department of Nuclear Medicine, Seoul National University Bundang Hospital, Seongnam, Korea
[7]Information Systems and Technology, Seoul Metropolitan Government-Seoul National University Hospital Boramae Medical Center, Seoul, Korea
[8]Center for Data Science, Biomedical Research Institute, Seoul Metropolitan Government-Seoul National University Hospital Boramae Medical Center, Seoul, Korea

Objectives: In the era of the Fourth Industrial Revolution, where an ecosystem is being developed to enhance the quality of healthcare services by applying information and communication technologies, systematic and sustainable data management is essential for medical institutions. In this study, we assessed the data management status and emerging concerns of three medical institutions, while also examining future directions for seamless data management. **Methods:** To evaluate the data management status, we examined data types, capacities, infrastructure, backup methods, and related organizations. We also discussed challenges, such as resource and infrastructure issues, problems related to government regulations, and considerations for future data management. **Results:** Hospitals are grappling with the increasing data storage space and a shortage of management personnel due to costs and project termination, which necessitates countermeasures and support. Data management regulations on the destruction or maintenance of medical records are needed, and institutional consideration for secondary utilization such as long-term treatment or research is required. Government-level guidelines for facilitating hospital data sharing and mobile patient services should be developed. Additionally, hospital executives at the organizational level need to make efforts to facilitate the clinical validation of artificial intelligence software. **Conclusions:** This analysis of the current status and emerging issues of data management reveals potential solutions and sets the stage for future organizational and policy directions. If medical big data is systematically managed, accumulated over time, and strategically monetized, it has the potential to create new value.

Keywords: Data Management, Big Data, Health Information Systems, Electronic Health Records, Medical Records

---

I. Introduction

With the advent of the Fourth Industrial Revolution, an ecosystem designed to enhance the quality of medical services is emerging in the healthcare sector, driven by the application of information and communication technologies (ICT) such as big data analysis and utilization, as well as artificial intelligence (AI). In South Korea, tertiary hospitals in particular are undergoing innovative digital transformations, highlighting the importance of managing the continually growing amount of healthcare information data.

As the utilization of medical data in public healthcare has become more prevalent, the government introduced the concept of pseudonymous information through the amendment of the Data 3 Act in 2020, laying the foundation for widespread use of medical data. South Korea has supported numerous healthcare information technology (IT) projects, including the Electronic Medical Record (EMR) certification program, a project improving the standardization of medical information, and pilot projects for the exchange of clinical information between hospitals [1]. Many ongoing initiatives support data-driven hospitals to promote research and foster the development of new medical technologies [2]. A project focused on individual-led medical data, known as the personal health record project, has been undertaken to empower individuals in managing their own medical records, life log health information, and public institution information through smartphone apps [3]. A pilot service called My-HealthWay was launched in conjunction with this initiative [4]. Additionally, a project established P-HIS, a cloud-based next-generation precision medicine hospital information system (HIS).

In order to keep up with current trends, such as building a medical big data-based platform and activating AI applications, managing clinical data accumulated in hospitals is crucial. Furthermore, effectively using infrastructure capable of sharing pseudonymized and standardized research data with researchers and industries is essential, particularly in medical treatment and research, which lie at the core of the information society. To this end, this study examined the current status of data management and emerging issues at Seoul National University Hospital (SNUH), SNU Bundang Hospital (SNUBH), and Seoul Metropolitan Government-Seoul National University Boramae Medical Center (SNU-BMC). The goal of this paper is to explore directions that promote seamless data management, organizational operation, and policy guidelines for the future.

II. Current Status of Medical Data Management

1. Data Types and Amount
HIS data are categorized into Electronic Medical Records (EMRs), picture archiving and communication systems (PACS), digital pathology (DP), clinical data warehouse (CDW), and research systems using a common data model (CDM). Table 1 shows the total amount of data in the HIS for each hospital.

2. Data Management Infrastructure
To store and manage data, hospitals choose on-premise or cloud-based approaches, or a combination of both to run their information systems. On-premise methods involve the hospital maintaining its own server, physically located on-site and managed by the IT department. Alternatively, hospitals can opt for a cloud system, in which an external cloud service provider supplies the necessary infrastructure support and security for managing data. A cloud system can be accessed through a subscription-based payment system, and resources and infrastructure can be easily expanded. The main benefit of a cloud system is reducing the initial budgetary burden for purchasing systems.

1) On-premise system
In 2004, SNUH implemented its first HIS, and in 2016, the HIS was upgraded to BestCare 2.0, a next-generation HIS. Following the plan to relocate the computer center from inside the hospital to an external location, the main comput-

Table 1. Total amount of data in hospital information systems (unit: GB)

Hospital	EMR	PACS	DP	CDW	CDM
Seoul National University Hospital	58,025	1,329,450	2,235,392	16,261	1,924
Seoul National University Bundang Hospital	12,887	613,606	3,497	5,501	405
Seoul Metropolitan Government–Seoul National University Boramae Medical Center	4,246	280,000	700 (DP System X)	601	92

EMR: electronic medical record, PACS: picture archiving and communication systems, DP: digital pathology, CDW: clinical data warehouse, CDM: common data model.

ers are currently situated in an external Internet data center (IDC), with the disaster recovery server (a simplified version of the HIS) installed in a separate space within SNUH (Figure 1). Both SNUBH and SNU-BMC have also installed and are operating HISs on their premises.

2) Cloud system

As a research system, SNUH employs Amazon Web Services to establish a CDM analysis environment and a big data analysis platform called SUPERB in a cloud environment. At SNUBH, a CDM is utilized both on-premises within the hospital and in a cloud system hosted by Amazon Web Services (Figure 2).

3. Data Backup Plans

SNUH conducts daily backups of the HIS database and server, managing all data in accordance with a policy that requires retention for three months. Additionally, SNUH performs a vault backup on hard disk drives once a month, retaining them for up to 2 years. Similarly, SNUBH categorizes the HIS data into three distinct types: EMR, PACS, and optical files. Primary backups are stored within the hospital, secondary backups are physically located in a separate location, and third backups are stored on hard disk drives and tapes.

4. Organizations for Data Management

(1) At SNUH, the data management organizational structure comprises the Information Planning Team, Information Development Team, Medical Records Team, and Information Security Center, all of which fall under the Office of Hospital Information. In addition to these teams, SNUH has the Department of Data Science Research, Department of Medical Device Innovation Research, and Department of Medical Device Development Support under the Biomedical Research Institute (Table 2). Other data management committees include the Healthcare Information Management Committee, which is responsible for evaluating the feasibility of projects related to information system establishment, and the Medical Records Management Committee, which oversees the quality of medical records and standardization of terminologies. The Institutional Review Board (IRB) reviews research projects that utilize medical data and the Big Data Review Board (BRB) is responsible for assessing the feasibility of utilizing big data.

(2) At SNUBH, the data management organizational structure includes the Office of Hospital Information, Digital Healthcare Research Department, Healthcare ICT Research Center, Big Data Center, and Medical AI Center (Table 3). When data is used for educational and research purposes, both the IRB and the Data Review Board work together to
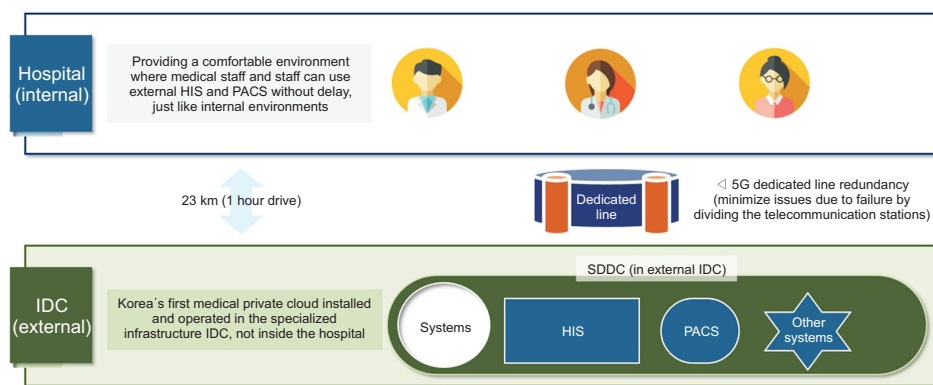

Figure 1. Schematic diagram of the computer center of Seoul National University Hospital. HIS, hospital information system; PACS, picture archiving and communication system; IDC, internet data center; SDDC, software-defined data center.
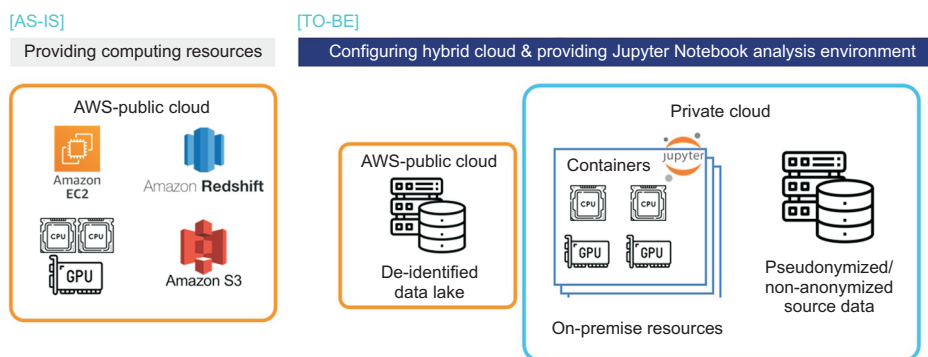

Figure 2. Big data analysis platform at Seoul National University Bundang Hospital. AWS: Amazon Web Services.

Table 2. Data management organizations of Seoul National University Hospital (SNUH)

Organization name	Key responsibilities
Office of Hospital Information	
Chief Information Officer and Deputy CIO	Overall management and decision-making
Information Planning Team	IT system project planning, execution, operation and maintenance, and outsourcing-related tasks, administrative tasks
Information Development Team	Development and operation of medical information systems, coordination and support of information-related tasks
Medical Records Team	Medical records management, disease-specific statistics processing, medical records research support, discharge patient analysis, cancer registry, record form development and terminology registration and management
Information Security Center	Establishment of information security and protection policies, planning, coordination, user education, public agency liaison, international certification, security audit tasks
Biomedical Research Institute: SNUH Innovative Medical Technology Research Institute	
Department of Data Science Research	Big data infrastructure support, data engineering, analytical statistics, AI research support
Department of Medical Device Innovation Research	Smart ICT, medical robotics and mechatronics, medical imaging devices, bio-materials/in-vitro diagnostics research labs
Department of Medical Device Development Support	Medical device licensing support, medical device clinical trial support, usability evaluation, prototype manufacturing support

COI: chief information officer, IT: information technology, ICT: information and communication technology, AI: artificial intelligence.

Table 3. Data management organizations of Seoul National University Bundang Hospital

Organization name	Key responsibilities
Office of Hospital Information	
Chief Information Officer and Faculty Members	Overall management and decision-making
Information Planning Team	Hospital information system planning and administrative tasks
Information Operations Team	Hospital information system operation and maintenance
Medical Records Team	Management of EMR, cancer registration, standard terminology, and electronic consent forms
Information Security Center	Establishment of information security policies and security certification, and management of personal information protection
Digital Healthcare Research Department	Domestic and international business support and external cooperation for hospital information systems
Healthcare ICT Research Center	Conducting external medical IT R&D and ICT convergence research projects, establishment and operation of big data and CDM systems
Big Data Center	Research and operational management in the big data field
Medical AI Center	Research and evaluation of the clinical utility of AI technology

EMR: electrical medical record, IT: information technology, ICT: information and communication technology, AI: artificial intelligence CDM: common data model.

ensure that data utilization is safe and responsible [5].

(3) At SNU-BMC, the data management organizational structure comprises the Healthcare Information Team, Medical Records Team, and EMR Operation Team under the Office of Hospital Information, as well as the Data Science Center under the Biomedical Research Institute (Table 4). SNU-BMC also operates both the IRB and the Data Review Board together for safe research.

III. Data Management Issues

1. Challenges of Data Management

1) Rapid data growth in data storage

In general, PACS capacity has shown an annual average increase of 11.25% at SNUH (Figure 3). Projections indicate that total storage will increase by approximately 1 petabyte (PB) per year from 2021 to 2025, about 0.5 PB per year from 2026 to 2028, and roughly 0.4 PB per year from 2029 onwards (Figure 4). The HIS, as well as the PACS, must be

Table 4. Data management organizations of Seoul Metropolitan Government-Seoul National University Boramae Medical Center

Organization name	Key responsibilities
Office of Hospital Information	
Chief Information Officer	Overall management and decision-making
Faculty Members	Medical information protection, EMR operation, PACS operation
Medical Information Team	Hospital information system planning and administrative tasks
Medical Records Team	Management of EMR, cancer registration, standard terminology, and electronic consent forms
EMR Operations Team	Medical information system operation and maintenance
Biomedical Research Institute	
Data Science Center	Management of data science research

EMR: electronic medical record, PACS: picture archiving and communication systems.
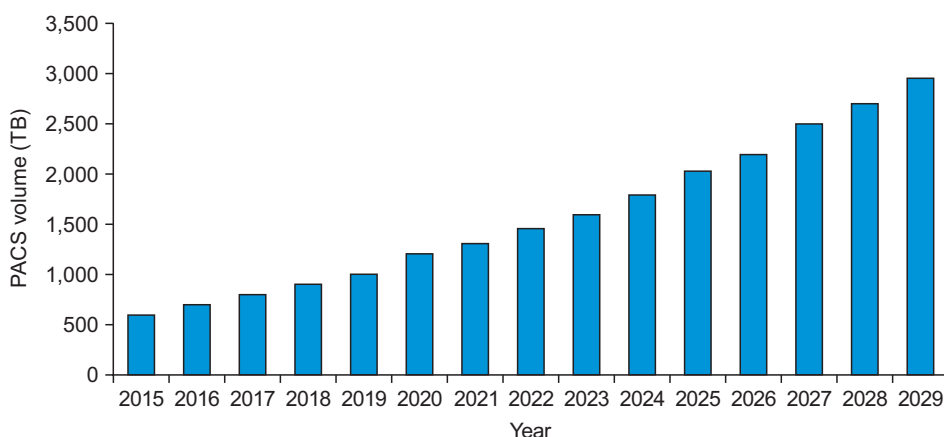

Figure 3. Picture archiving and communication systems (PACS) volume growth at Seoul National University Hospital.
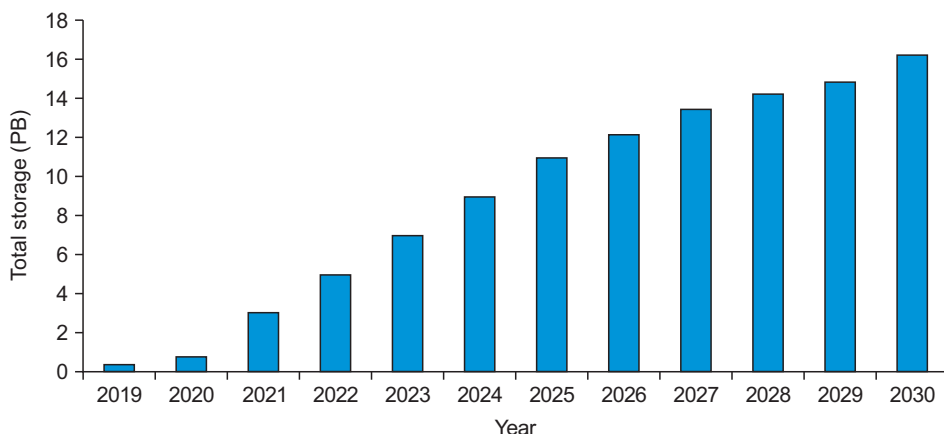

Figure 4. Storage growth over time.

able to handle continuous growth in users, patients, and data while maintaining performance and efficiency. However, as the demand and cost for data storage space continue to rise, maintaining all data in the operating system may become unsustainable, considering the significant system purchase costs and management manpower expenses required. To achieve cost and manpower efficiency, it is recommended to separate clinical operational data and research data in the future to keep the speed-critical clinical database lightweight.

2) The need for research resource management

There are management issues related to the cost of continuous storage expansion and monitoring data collection infrastructure while gathering real-time bio-signal data for research purposes. Moreover, when converting CDM data annually, problems arise with previous CDM data, as multiple versions of the CDM database must be maintained for at least 4 or 5 years to publish a research paper.

Insufficient management after project completion poses challenges for hospitals in disposing of or redistributing hardware resources, such as servers, acquired by in-house researchers through research projects. Consequently, research systems not included in the main system create significant security loopholes, as management manpower vanishes or management levels decline once tasks are completed. Thus, it is crucial to establish a stable management protocol capable of effectively managing all resources within the computer center.

3) Labor cost of data management

Insufficient support for information organizational structures (particularly regarding labor costs) in hospitals is a widespread problem, not exclusive to SNUH. This occurs because hospitals often prioritize profit over investing in infrastructure, which can lead to poor data management.

In some hospitals, the responsibility for server and research data management is primarily assigned to personnel hired through research funds. This can create challenges in maintaining consistent and reliable data management practices in the future.

2. Government Regulations

Regarding government support and regulation, the following issues need to be considered.

Regulations on the "Retention of medical records data" are detailed in Article 15 of the Enforcement Rule of the Medical Service Act (Preservation of Medical Records) [6] as shown in Figure 5.

The Medical Service Act and its Enforcement Rule systematically outline the retention period for medical records. However, the Medical Service Act does not provide separate regulations regarding the management or destruction of medical records beyond what is stipulated in the Personal Information Protection Act. This lack of clear guidance is further complicated by the ambiguous explanation of the "proviso clause" in Article 15, Paragraph 1 of the Enforcement Rule of the Medical Service Act. As a result, effectively managing data and systems in the field becomes difficult. Medical institutions, including SNUH and SNU-BMC, at which HISs have been in operation for more than 10 years, must decide whether to retain or destroy medical records in compliance with applicable regulations. However, detailed destruction guidelines and procedures are absent at this moment. Furthermore, the regulation mandating data "destruction" after the retention period has expired is based on personal information protection and does not account for other secondary uses such as long-term treatment or research. Consequently, there is a need for institutional guidelines addressing the retention, destruction, and management of various types of data in medical institutions.

In order to vitalize the digital healthcare industry, some areas need improvement in terms of health and medical data integration and utilization. A medical data-driven hospital project [2] is being carried out to exchange standardized medical data, but only data exchange between institutions within the consortium is currently possible. Through the MyHealthWay project [4], it is necessary to establish an ecosystem that integrates individual-centric medical data, allowing individuals to transmit and utilize their own data. With the amendment of the Data 3 Act, data for research can be provided after pseudonymization, but conflicts with

---

① Medical personnel or founders of medical institutions shall preserve medical records under Article 22, Paragraph 2 of the Medical Service Act for the period specified in each of the following subparagraphs.
1. Patient list: 5 years,
2. Medical records: 10 years,
...
9. Copy of medical certificate: 3 years
In addition to the subparagraphs of Article 15, Paragraph 1, the following proviso has been newly established:
However, if continuing medical treatment is necessary, the period may be extended and preserved within the period specified in each of the following subparagraphs, limited to one time.

Figure 5. Details of Article 15 of the Enforcement Rule of the Medical Service Act (preservation of medical records).

the existing Medical Service Act [7], which prohibits the disclosure of medical records to third parties, must be resolved. Additionally, the right to request transfer of medical data to third parties should be strengthened through legislation. To regulate AI-based medical devices in Korea, the Ministry of Food and Drug Safety distributes and continuously revises guidelines for the approval and clinical efficacy evaluation of medical devices that incorporate big data and AI technology [8-10]. Furthermore, as generative AI has recently attracted attention, guidelines for trustworthy and explainable AI are also needed.

3. Need for Comprehensive Metadata Management

Effective metadata management is essential for comprehending the meaning and relationships within complex healthcare big data and enhancing data transparency. By minimizing confusion and errors resulting from unclear data, metadata management enables quick understanding and navigation of the data. This is crucial for data quality control and improvement initiatives, facilitating accurate decision-making and swift access to new insights. Therefore, to improve the knowledge and utilization of medical big data, a systematic and ongoing protocol for metadata management should be established throughout the entire hospital organization and operated accordingly. SNU hospitals use the same terminology sets for diagnosis, chief complaints, and operations through the standard protocol for adopting new terms. In the future, collaborative work across hospitals will be necessary to expand terminology standardization for other areas such as nursing and clinical lab tests. Effective metadata management, backed by standard terminology, can substantially enhance data consistency, interoperability, and quality within SNU hospitals. A comparative analysis would help highlight current strengths, identify areas for improvement, and guide a roadmap for future enhancements, including potential inter-institutional cooperation to harmonize data practices across different hospitals.

IV. Further Considerations for Data Management

1. Promoting Data Sharing among Hospitals

Fast Healthcare Interoperability Resource (FHIR) is a standard framework for healthcare information exchange created by HL7 (Health Level 7). In Korea, the adoption of the FHIR standard is being promoted through initiatives led by the Korea Health Information Service and HL7 Korea, aiming to support online education and technology [11]. The SNUH,

SNUBH, Gangnam Center, and SNU-BMC developed a standard terminology management system in 2013 and integrated it into their HIS. SNUBH plans to expand and incorporate the HL7 FHIR standard, providing a foundation for patient-centered data sharing, enhancing Health4U (a personal health record portal/app for patients) and building an infrastructure capable of delivering patient-generated health data to EMRs. At SNU-BMC, adopting the HL7 FHIR next-generation standard framework is vital, and there is also a need to consider offering mobile patient services for access via mobile devices and web portals through the standard FHIR API. If government-level guidelines are introduced in this area, an efficient system can be established.

2. Infrastructure Configuration for AI Applications

To clinically evaluate AI software developed as a research project or to integrate a solution in the process of commercialization, it is crucial to install and assess the application without disrupting the existing EMR system. Creating an environment that accommodates AI technology necessitates a thorough review of patient safety issues and ethical considerations. A digital twin of real-world data, including people, hospitals, systems, and diseases, can enable safe and effective simulations in virtual reality [12]. These simulations can enable the development of various applications such as training, education, remote medical consultations, and medical AI services. However, there is currently no dedicated department in the hospital, and implementing such an environment depends on the determination of hospital administrators to address these conditions.

V. Discussion

In order to foster innovation that converges medicine and technology through the digital transformation of medical institutions, systematic and continuous data management through data-related organizations must be executed. To this end, the current status of SNUH, SNUBH, and SNU-BMC was investigated by data type, management infrastructure, backup plans, and management/organizational structure. Issues arising during data management include cost concerns due to increased storage capacity, an insufficient resource management system, problems with government data management regulations, the need for enterprise-wide metadata management, and personnel retention issues. To address these problems, each hospital needs to recognize the importance of data management, provide appropriate support, and share and discuss problem solutions among hospitals.

Additionally, government support for data management and identification of realistic problems with data storage and institutional improvements are necessary.

Data are swiftly becoming a valuable asset in society. To effectively manage the vast amount of data retained by each hospital, it is essential to progressively expand standards and examine the direction and system of providing mobile patient services. Furthermore, there is a need for a platform that can verify and apply various AI applications.

To prepare for the rapidly changing world, hospitals need to introduce new technologies such as artificial intelligence and the metaverse. Managing the accumulated medical big data is a shortcut to introduce these technologies and create new value, preparing well for the future.

Conflict of Interest

Jinwook Choi is an editorial member of Healthcare Informatics Research; however, he did not involve in the peer reviewer selection, evaluation, and decision process of this article. Otherwise, no potential conflict of interest relevant to this article was reported.

ORCID

Jinwook Choi (https://orcid.org/0000-0002-9424-9944)
Hyeryun Park (https://orcid.org/0000-0002-0923-8491)
Eui Kyu Chie (https://orcid.org/0000-0003-2027-7472)
Sae Won Choi (https://orcid.org/0000-0002-0123-8227)
Ho-Young Lee (https://orcid.org/0000-0001-6518-0602)
Sooyoung Yoo (https://orcid.org/0000-0001-8620-4925)
Byoung Jae Kim (https://orcid.org/0000-0002-2734-997X)
Borim Ryu (https://orcid.org/0000-0002-2000-4565)

References

1. Korea Health Information Service. Standardization of health and medical information [Internet]. Seoul, Korea: Korea Health Information Service; c2023 [cited at 2023 Aug 2]. Available from: https://www.k-his.or.kr/menu.es?mid=a10203040000.
2. Korea Health Information Service. Medical data-driven hospital portal [Internet]. Seoul, Korea: Korea Health Information Service; c2023 [cited at 2023 Aug 2]. Available from: https://www.hins.or.kr/index.es?sid=a2.
3. Korea Health Information Service. Personal health record (PHR) [Internet]. Seoul, Korea: Korea Health Information Service; c2023 [cited at 2023 Aug 2]. Available from: https://www.k-his.or.kr/menu.es?mid=a10204000000.
4. Cho W. Trial opening of medical MyData, My HealthWay [Internet]. Seoul, Korea: MediGateNews; 2022 [cited at 2023 Aug 2]. Available from: https://medigatenews.com/news/2275592704.
5. Seoul National University Bundang Hospital. Data review board [Internet]. Seongnam, Korea: Seoul National University Bundang Hospital; c2023 [cited at 2023 Aug 2]. Available from: https://www.snubh.org/member/deliberation.do.
6. Korea Law Information Center. Enforcement Rule of the Medical Service Act [Internet]. Sejong, Korea: Ministry of Government Legislation; 2015 [updated at 2022 Nov 22, cited at 2023 Aug 2]. Available from: https://law.go.kr/법령/의료법%20시행규칙/제15조.
7. Korea Law Information Center. Medical Service Act [Internet]. Sejong, Korea: Ministry of Government Legislation; 2002 [updated at 2020 Dec 29, cited at 2023 Aug 2]. Available from: https://glaw.scourt.go.kr/wsjo/lawod/sjo192.do?lawodNm=의료법&jomunNo=21&jomunGajiNo=2.
8. Ministry of Food and Drug Safety. Guidelines for approval review of medical devices utilizing big data and artificial intelligence technology [Internet]. Cheongju, Korea: Ministry of Food and Drug Safety; 2019 [cited at 2023 Aug 2]. Available from: https://www.mfds.go.kr/brd/m_1060/view.do?seq=14407&srchFr=&srchTo=&srchWord=&srchTp=0&itm_seq_1=0&itm_seq_2=0&multi_itm_seq=0&company_cd=&company_nm=&Data_stts_gubun=C1004&page=2.
9. Ministry of Food and Drug Safety. Guidelines for clinical efficacy evaluation of artificial intelligence-based medical devices [Internet]. Cheongju, Korea: Ministry of Food and Drug Safety; 2019 [cited at 2023 Aug 2]. Available from: https://www.mfds.go.kr/brd/m_1060/view.do?seq=14406&srchFr=&srchTo=&srchWord=&srchTp=0&itm_seq_1=0&itm_seq_2=0&multi_itm_seq=0&company_cd=&company_nm=&Data_stts_gubun=C1004&page=2.
10. Ministry of Food and Drug Safety. Guidelines for approval and review of artificial intelligence medical devices [Internet]. Cheongju, Korea: Ministry of Food and Drug Safety; 2022 [cited at 2023 Aug 2]. Available from:

https://www.mfds.go.kr/brd/m_1060/view.do?seq=1501
1&srchFr=&srchTo=&srchWord.

11. Korea Health Information Service. Healthcare information standardization [Internet]. Seoul, Korea: Korea Health Information Service; c2023 [cited at 2023 Aug 2]. Available from: https://www.k-his.or.kr/menu.es?mid=a20203040000.

12. Korea Health Information Service. Innovation of healthcare utilization data (Brief 2022 Vol.3) [Internet]. Seoul, Korea: Korea Health Information Service; 2022 [cited at 2023 Aug 2]. Available from: https://www.k-his.or.kr/board.es?mid=a10309000000&bid=0025.