Review Article

Healthc Inform Res. 2019 October;25(4):239-247.
https://doi.org/10.4258/hir.2019.25.4.239
pISSN 2093-3681 • eISSN 2093-369X

HIR
Healthcare Informatics Research

Protecting and Utilizing Health and Medical Big Data: Policy Perspectives from Korea

Dongjin Lee, Mijeong Park, Seungwon Chang, Haksoo Ko
School of Law, Seoul National University, Seoul, Korea

Objectives: We analyzed Korea's data privacy regime in the context of protecting and utilizing health and medical big data and tried to draw policy implications from the analyses. **Methods:** We conducted comparative analyses of the legal and regulatory environments governing health and medical big data with a view to drawing policy implications for Korea. The legal and regulatory regimes considered include the following: the European Union, the United Kingdom, France, the United States, and Japan. We reviewed relevant statutory materials as well as various non-statutory materials and guidelines issued by public authorities. Where available, we also examined policy measures implemented by government agencies. **Results:** In this study, we investigated how various jurisdictions deal with legal and regulatory issues that may arise from the use of health and medical information with regard to the protection of data subjects' rights and the protection of personal information. We compared and analyzed various forms of legislation in various jurisdictions and also considered technical methods, such as de-identification. The main findings include the following: there is a need to streamline the relationship between the general data privacy regime and the regulatory regime governing health and medical big data; the regulatory and institutional structure for data governance should be more clearly delineated; and regulation should encourage the development of suitable methodologies for the de-identification of data and, in doing so, a principle-based and risk-based approach should be taken. **Conclusions:** Following our comparative legal analyses, implications were drawn. The main conclusion is that the relationship between the legal requirements imposed for purposes of personal information protection and the regulatory requirements governing the use of health and medical data is complicated and multi-faceted and, as such, their relationship should be more clearly streamlined and delineated.

Keywords: Big Data, De-identification, Data Protection, Privacy, Research

Submitted: August 15, 2019
Revised: October 6, 2019
Accepted: October 13, 2019

Corresponding Author
Haksoo Ko
School of Law, Seoul National University School of Law, 1, Gwanak-ro, Gwanak-gu, Seoul 08826, Korea. Tel: +82-2-880-2602, E-mail: hsk@snu.ac.kr (https://orcid.org/0000-0002-6968-774X)

I. Introduction

Big data in the health and medical area has great potential for utilization. Analyses of health and medical big data may bring about exceedingly useful and fruitful results. For instance, precision medicine, which requires data on individuals from various sources, is a promising field with great potential for widespread adoption and application in the future [1].

Yielding results that can usefully be adopted often requires, as a pre-requisite, the compilation of a large amount of data from diverse sources. Compiling a large amount of data, however, could serve as a double-edged sword. On the one hand, this may lead to enhanced capability for data analytics and to improvements in diagnostics and treatments. On the other hand, as a general matter, there could be a concern regarding data privacy risks associated with a large amount of compiled health and medical data. Some argue that the current legal regime for personal data protection has yet to come up with a satisfactory solution regarding the difficult question as to how we may offer an appropriate level of data privacy protection while taking advantage of the potential that can be realized by effectively utilizing big data.

In the following sections, we present comparative analyses of the legal and regulatory environments governing health and medical big data with a view to drawing policy implications for Korea. The legal and regulatory regimes considered include the European Union (EU), the United Kingdom (UK), France, the United States (US), and Japan. We review relevant statutory materials as well as various non-statutory materials and guidelines issued by public authorities. We then summarize implications from the comparative analyses.

II. Health and Medical Information: Defining Relevant Terms

In the realm of health and medical big data, in terms of legal terminology employed in Korea, there are several important terms that should be taken into consideration. They include such terms as genetic, biometric, and sensitive information. Definitions of these concepts are not always very clear, and the relationships among these terms are often exceedingly complicated. First, about health and medical information in general, the Framework Act on Health and Medical Services specifically introduces the concept of "health and medical information" (Article 3(6)). This provision defines health and medical information as information that should be made widely available for public health purposes and provides illustrative examples, including "knowledge relating to healthcare, or all types of healthcare data which are expressed in symbols, numbers, letters, voice, sound, video, and other forms" [2]. This definition includes certain types of personal data, which could fall under the jurisdiction of laws governing data privacy. This approach of defining health and medical information is different from the approaches that can be found in some other jurisdictions. For instance, in the EU, the General Data Protection Regulation (GDPR) defines health and medical information as part of special categories of personal data.

About genetic information, the Bioethics and Safety Act defines genetic information as information regarding the genetic characteristics of an individual, which is obtained by analyzing human body components or biospecimens (Article 2, Subparagraphs 11 and 14). While this definition aims at delineating genomics research and at addressing concerns over the possibility of discrimination, it also deals with privacy aspects. Another statutory provision on genetic information can be found in the Personal Information Protection Act (PIPA), which is a general statute dealing with issues of data privacy. Its Enforcement Decree explicitly provides that genetic information is part of the statutorily defined sensitive information (Article 19, Subparagraphs 1 and 2). This way of defining genetic information is similar to the approach taken in the EU GDPR, which, as noted above, defines special categories of personal data as a separate type of personal data. Meanwhile, in Japan, DNA sequence information is defined as part of the "personal data requiring special care" under the Act on the Protection of Personal Information. In the case of the US, the Genetic Information Nondiscrimination Act (GINA) declares that genetic information requires protection. Thus, it appears that, in many jurisdictions including Korea, genetic information is considered a part of personal data, requiring an enhanced level of protection. This may reflect certain unique characteristics of genetic information. That is, genetic information is in general invariant; members of the same family share much of the same genetic sequence; and an individual can often be identified using genomic sequence data.

Separate from genetic information, biometric information could be considered a separate category of personal information. In Korea, there is a separate statutory provision defining biometric information. The Act on the Promotion of Information and Communications Network Utilization and Data Protection (IC Network Act) defines biometric information as "the information related to the physical or behavioral characteristics of an individual which may be used for identification and which may include the information such as fingerprints, iris data, voice data, and handwriting samples" (IC Network Act, Enforcement Decree, Article 9-2, Paragraph 1, Subparagraph). In terms of data privacy, the main concern over the use of biometric information is the possibility of identification or of individuation through biometric information. That is, there is a concern that certain types of biometric information could be used as an identifier in a dataset. At the same time, there is a growing use of wearable and other devices that are equipped with biometric sensors, and questions are being raised as to whether such biometric sensor data should be considered personal information and/or health information.

Under Korea's PIPA, it is not entirely clear, but it could be argued that medical information, genetic information, and biometric information should all be deemed to be part of the statutorily defined sensitive information. In other jurisdictions, a similar statutory treatment can be found. Notably, the EU GDPR expressly stipulates that special categories of personal data include 'genetic data', 'biometric data for the purpose of uniquely identifying a natural person', and 'data concerning health' (Article 9.1). In the US, there are statutory and other provisions that consider health records to be sensitive personal information [3]. In Japan, the concept of the 'information requiring special care' is important, and under Japanese law, this category of personal information includes certain medical information that may give rise to concerns over "unfair discrimination, prejudice, or any other disadvantage to an individual". In Korea's PIPA, the definition of 'sensitive information' contains an illustrative list, and this list does not include medical information (Article 23, Paragraph 1). This provision, however, includes 'information concerning health'. The conceptual distinction between health information and medical information and their relationship have not been clearly resolved [4].

With the backdrop on legal definitions of health and medical information and other related concepts, we now consider what specific legal and regulatory issues that are being discussed in Korea. First, it is being debated whether health and medical information may be processed for scientific research purposes, statistical purposes, or other purposes, without having to obtain consent from data subjects. A similar question about data use is being raised about using genetic, biometric, and other related information. Second, it is being debated whether (and how) it would be possible to achieve the dual goal of (1) providing adequate protection to personal information, while (2) fostering active research and development activities in the healthcare and medical sector.

Third, related to the above, there are various legal and regulatory questions being raised. They include, for instance, debates on the precise meaning of 'identification' in the medical and healthcare context, on the proper methodology and procedure for de-identification (including pseudonymization) of health and medical information; and on the use of health and medical information for secondary purposes or for purposes that are related but arguably different from the original purpose given at the time of initial collection. Fourth, there are also debates as to whether it is necessary to obtain consent and how to safeguard personal information in the context of clinical research and clinical trials.

III. A Comparative Legal and Regulatory Overview

1. European Union

In May 2018, the EU began implementing the GDPR, which represents a unified EU-wide legal and regulatory scheme on personal data protection. A notable characteristic of the GDPR is that it contains several explicit provisions that are related to health and medical data. Perhaps most significantly, while the GDPR stipulates that health data falls under special categories of personal data, it also provides certain exceptions under which health data may be utilized without obtaining consent from data subjects, for instance, for scientific research purposes, provided that requisite safeguards are in place. Separately, in March 2019, the Council of Europe finalized and published its Recommendation on the Protection of Health-Related Data, which offers guidance regarding the processing of personal health data [5]. There is also a directive on the re-use of public sector information [6]. This directive allows for the re-use of information for commercial purposes as well as for non-commercial purposes if such information is generally accessible to the public. This opens certain avenues for utilizing data for a secondary purpose in the context of clinical trials.

2. United Kingdom

The Data Protection Act of 2018 serves as a general statute governing data protection in the UK. As a general matter, the Data Protection Act is very similar to the GDPR. In the realm of healthcare, the Health and Social Care Act of 2012 is important: this Act enables integration of healthcare and social security services that had been carried out by the NHS. Also, through this Act, NHS Digital was established, and attempts have been made to streamline data gathering and sharing processes.

Meanwhile, the UK established a National Data Guardian, which provides guidance relating to the protection of the health and medical information of its citizens, while allowing for appropriate data utilization. A series of reports on health and medical data, entitled Caldicott Reports, have been published in this vein, promulgating fundamental principles to be employed. Notably, recent Caldicott Reports specifically emphasized the significance of effectively utilizing data in addition to properly protecting data [7].

Finally, the UK established a biobank that collects various data, including biometric information, environmental factors, past and present health conditions, and other physical characteristics for roughly 500,000 UK residents between the ages of 45 and 69. Additionally, the UK 100,000 Genome Projects aim at compiling genetic sequencing data for 100,000 individuals [8,9].

3. France

In France, the Law on Information Processing, Documents, and Liberties (La loi no 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés) regulates the protection of personal data [10], while the recently enacted Law for a Digital Republic (Loi pour une République numérique) covers such areas as net neutrality, data transfer, and the right to remain connected. A notable feature of these laws is that, in addition to public data and private data, they designate a third category of data and require the data to be made public. Also, the Law on the Modernization of Our Health System (Loi de modernisation de notre système de santé) established simplified processes for obtaining consent for research purposes and, at the same time, reinforced the criteria for liability. Based on this law, the National Healthcare Insurance Fund established a national system for health data (système national des données de santé). Further, the rules for providing health information were codified in the Code of Public Health (Code de la santé publique).

Meanwhile, the Law on Bioethics (Loi de la Bioéthique) contains provisions on the necessary criteria to be used for obtaining consent in relation to the examination of genetic characteristics (examen des caractéristiques génétiques), which allows for secondary use under certain conditions.

4. United States

In the US, perhaps the most important statute that governs health information would be the Health Insurance Portability and Accountability Act (HIPAA). The HIPAA Privacy Rule, a subsidiary rule to the HIPAA, establishes a national standard on the protection of personal health records and health information [11]. The HIPAA Privacy Rule also contains detailed provisions on the de-identification of personal health information. The HIPAA Security Rule establishes general safety and security conditions for the protection of health information.

Separately, the Health Information Technology for Economic and Clinical Health Act (HITECH Act) contains Subtitle D, which covers security issues related to electronic health records. Further, the HITECH Act provides for financial incentives to healthcare organizations for adopting a system for electronic health records [12]. In 2016, the 21st Century Cure Act was enacted, which contains provisions aimed at facilitating data sharing to enhance the effectiveness of patient care by encouraging the development of new drugs and other medical products.

Rules of ethics for federally funded research, which are often referred to as the Common Rule, also play an important role in the context of utilizing healthcare information. The Common Rule is a set of ethical rules on human subject research, and it is codified in the form of federal regulations. A recent amendment to the Common Rule introduced the concept of ‘broad consent’ and made the use of a single Institutional Review Board (IRB) mandatory.

Recently, genomic information also has received much attention. In particular, the National Institute of Health created the Genomics Network, with the goal of facilitating the collection of biospecimens. Access to federal databases has been made available through HealthData.gov, which allows for the downloading of certain types of data.

5. Japan

Japan has recently amended its data privacy statute, the Act on the Protection of Personal Information [13]. Through the amendment, the concept of personal information was clarified, and a new concept of ‘anonymously processed information’ was introduced. Also, a separate statute was promulgated to cover health and medical information. That is, the newly enacted Act on Anonymously Processed Medical Information to Contribute to Medical Research and Development of Japan aims at fostering the utilization of anonymously processed data for research and development purposes. The same Act explicitly refers to the term of ‘anonymously processed information’, and it allows for the use of anonymously processed information with an opt-out mechanism. In addition, the Act to Promote Healthcare and Medical Strategy offers guidance regarding genetic testing and diagnosis [14]. In terms of governance, the Headquarters for Health Policy, established in 2013 under the auspices of the Prime Minister’s Office, is mainly responsible for crafting the agenda for health and medical policy.

6. Korea

The PIPA, enacted in 2011, serves as a general statute covering data privacy issues in Korea. The PIPA defines personal information as information that enables, directly or when ‘easily combined with other information’, the identification of individuals (Article 2). Thus, identification through a combination of information from a few different sources could possibly satisfy this statutory definition of personal information. There have been a lot of debates surrounding the precise concept of and criteria for identification, and further discussions were made as to how to ‘de-identify’ personal information. There are unresolved debates about the claims that most, if not all, de-identified personal information can be re-identified.

Personal information should, in most situations, be processed pursuant to the original purpose explained to the data subject at the time of data collection and with the consent of the data subject. In certain limited circumstances, personal information may be processed for purposes other than the initial purpose that was explained to the data subject, as stipulated by applicable laws and regulations. Also, regarding ‘sensitive information’, which includes health information, a separate consent should be obtained. Resulting practical difficulties from these requirements include the following: (1) obtaining consent for secondary purposes and (2) obtaining consent for the purposes of future research. Difficulties arise due to, among others, the fact that it could be hard to specify in advance what these purposes are and that these purposes could change as relevant circumstances evolve.

Separately, the Bioethics and Safety Act is often relevant in the context of utilizing health and medical information. In particular, pursuant to this Act, biospecimen research becomes permissible only when a research plan is approved by an IRB and with written consent from biospecimen donors. Meanwhile, in the public domain, the Act on the Promotion of the Provision and Use of Public Data serves an important role. Pursuant to this Act, the National Health Insurance Service (NHIS) and the Health Insurance Review & Assessment Service (HIRA) provide access to the data that they hold for research purposes. Additionally, the National Evidence-Based Healthcare Collaborating Agency has the legal authority to make requests to government agencies and public organizations to submit certain health-related data, and these agencies and organizations are obliged to submit such data after redacting personally identifiable information.

Regarding the maintenance of medical records, while it has been mandatory to keep these records within the premises of individual hospitals and clinics, relevant provisions of the Medical Services Act were amended in 2016, and now medical records can be stored offsite. Through this amendment, employing cloud computing services became permissible.

IV. Research Activities and Technological Developments

Various attempts have been made in various jurisdictions to leverage the potential value of health and medical big data from technological and policy perspectives. Some of these attempts include the following.

1. Jurisdictions Outside Korea

The EU developed the Horizon 2020 Framework Program, which is the largest research support program in the history of the EU [15]. The Digital Single Market initiative also emphasizes the necessity of digital technology for health management and care. Separately, the AEGLE Project publishes reports on the current state of big data research [16].

In the UK, in 2018, the NHS issued recommendations regarding how NHS data can be shared with private companies and third parties [17]. NHS Digital also makes part of its datasets available through its website [18]. Regarding the de-identification of personal data, the Information Commissioner’s Office (ICO) published a code of conduct on anonymization in 2012, which has served as useful guidance for practitioners.

France has established a National System of Health Data (Système National des Données de Santé) and allowed certain types of health data to be used for research purposes with a simple notification to the government, provided that such data falls under the exceptions stipulated in Article 8, Section 2 of the Law on Information Processing, Documents, and Liberties (La loi no 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés). Also, the Epidemiologie-France portal offers access to approximately 260 public databases on healthcare and medicine as well as to 500 other databases [19]. In January 2018, the Commission nationale de l’informatique et des libertés (CNIL) published a new guideline on personal information protection, which systematically implements various measures requested by the GDPR [20]. The CNIL also announced conditions that commercial partners (partenaires commerciaux) need to comply with to legally share data with a third party.

In 2016, the US government announced the Open-Government Plan, which reflects an open-data policy and would facilitate data utilization [21]. Health-related data is offered separately through the HealthData.gov website. In total, almost 3,000 datasets are offered through this website, and they include data from most institutions holding public health and medical data as well as from state governments. Regarding data de-identification, the HIPAA Privacy Rule provides guidance regarding (1) expert determination and (2) safe harbor. The Department of Health and Human Services (HHS) also published a guideline on de-identifying health information [22].

Japan has allowed the creation of a big data exchange platform since October 2018, although it is unclear yet whether actual exchanges of healthcare data take place on a regular basis. The Personal Information Protection Commission has also issued a guideline related to 'anonymously processed information.'

2. Korea

In Korea, the NHIS manages health insurance big data. The NHIS maintains massive databases, covering virtually all residents in Korea, and it has built an additional database system for research purposes. Databases in this category include cohort data as well as certain 'customized data'. The NHIS maintains a procedure for giving access to these databases. The HIRA also provides access to data to researchers through its health and medical big data open system. Separately, the National Cancer Center maintains a national datacenter for cancer and conducts research activities regarding, among others, establishing a big data platform for precision medicine for cancer. Also, the Korea Centers for Disease Control and Prevention maintains a human bio-resource bank as part of a national endeavor to build and maintain biospecimen banks and resource centers for pathogens, while conducting the Korean Genome and Epidemiology Study at the same time.

Meanwhile, several large hospitals maintain their own big data centers. Some of them have developed internal procedures and systems for de-identifying medical data and for giving access to these data for research purposes. Separately, certain research-oriented hospitals strive to construct a platform for conducting research on certain specific diseases and specialized areas.

V. Implications from Comparative Analyses

From comparative analyses of relevant laws and regulations, the following implications can be drawn.

1. Legal and Regulatory Approaches

First, the relationship between (1) the legal requirements imposed for purposes of personal information protection and (2) the regulatory requirements governing the use of health and medical data is complicated and multi-faceted. Many countries appear to have introduced a separate legal regime exclusively applicable to health and medical data. This is different from the approach that Korea has developed and maintained so far, in which, among others, (1) health-related information is enumerated as part of 'sensitive information' under the PIPA, general data privacy law; (2) the legal definition of medical information is not very clear; and (3) data privacy aspects related to genetic information is only accorded partially in the context of biospecimen research under the Bioethics and Safety Act.

Further, while certain jurisdictions, notably the EU, provide a mechanism under which personal health and medical data can be used for research and other purposes, doing so would be exceedingly cumbersome in Korea. More specifically, under the GDPR, pseudonymized health and medical data can be used for scientific research purposes and for statistical purposes. Under Korean law, on the other hand, health information belongs to statutorily defined sensitive information, and explicit consent is required from data subjects before sensitive information can be utilized, unless a separate statutory ground exists obviating the need to obtain consent.

Separately, there is a need to examine existing approaches for the protection of health and medical information more closely. Utilizing healthcare data could sometimes entail sharing and combining data from different sources, and such sharing and combining may increase privacy risks. Thus, while it would be imperative to recognize the need for sharing and combining data under certain circumstances, appropriate safeguards should be established and put in place at the same time. These safeguards could include, for example, setting up procedures for granting access to data and for monitoring how such data is being used in actuality.

Also, there is a need to increase efforts towards standardization and the reinforcement of inter-operability. Adopting appropriate standards and enhancing inter-operability are crucial pre-requisites for data utilization. A certification mechanism for electronic medical record systems could be a start. More importantly, careful consideration should be given to the legal framework on information gathering and sharing as well as on the establishment of a proper institutional structure for coordination and monitoring activities. Detailed rules and procedures regarding data pseudonymization should be developed as well.

2. Data Governance and Supervision

When it comes to the utilization of health and medical big data, engaging diverse groups of stakeholders and paying enough attention to issues in data governance are crucial. Various interested groups, including patients, doctors, and other medical professionals, and hospitals, need to be consulted in the process of making important social decisions on health and medical big data. Transparency and accountability are crucial, and as such, it is critical to elicit consensus and voluntary participation through sufficient communication among these parties and important stakeholders. An institution or organization may need to be established to prepare and execute national strategies on health and medical big data.

An important question that can be raised in the process of establishing a governance structure for health and medical big data is how an institution or organization regarding, for instance, data sharing could be developed. In Korea, first of all, massive data is accumulated and stored in NHIS and HIRA databases. A harmonized and coherent procedure and methodology could be devised to conduct data de-identification and to give access to the data that these organizations hold to researchers. Separately, a mandatory disclosure system could be devised for publicly funded research activities. That is, a requirement could be introduced, under which data disclosure could in principle become mandatory for publicly funded research projects. In the long run, a model would need to be explored, under which patients are given more control, and data access can be granted based on such patients' control and consent. Developing such models would be especially important in cases where implementing appropriate de-identification measures may be difficult or impracticable. Genetic data could be a good example of such a case.

In terms of legal supervision and compliance, issues related to data access and sharing typically fall under the domains not just of data privacy but also of bioethics. Regulations in these domains reflect various sets of legislative and policy goals. Special attention needs to be paid to streamline the applicable procedures and principles.

3. Technical Measures

An important characteristic of health and medical information is that, in general, it is accumulated as time progresses and the relevant databases become larger. That is, information on individual data subjects typically accumulates as time goes by. Thus, de-identification may sometimes be ineffective since data that is de-identified at one point of time can become re-identifiable at a future point of time. Also, compared to other types of data, there are many features and characteristics that are unique to health and medical information. Thus, there is a need to develop de-identification mechanics and procedures that are tailored to reflect special circumstances associated with health and medical information. A caveat, however, would be that this would not mean providing detailed and rigid technical specifications. Instead, a principle-based and risk-based approach needs to be taken, and there should be enough room for flexibility when these mechanics and procedures are implemented. That way, future advancements of relevant technologies can easily be accommodated.

VI. Policy Implications

Based on the foregoing, the following policy implications can be drawn.

First, there is room for improvements in terms of the regulatory and compliance regime relating to health and medical data. As noted, there is a regulatory scheme based on a set of statutes governing data privacy in general, including health data. Debates have been made as to whether general data protection authorities have the requisite expertise and whether they are well prepared to handle data in the health and medical sector. At the same, when human subject research is conducted, an additional layer of protection is offered in the form of IRB reviews. During the IRB review process, data privacy issues are expected to be examined as well.

It is, however, unclear if the IRB review process is adequate in terms of providing assurance regarding data privacy. Perhaps reflecting concerns over IRB's expertise in data privacy issues, some large hospitals in Korea have established a separate review board to examine and supervise data use. While these layers may afford additional safety, they may also lead to concerns of overlapping and repetitive (and possibly ineffective) regulations. Further efforts are needed to streamline the overall process to enhance efficiency and effectiveness at the same time.

Second, to improve the environment for scientific research utilizing health and medical big data, the overall IRB review system may need to be re-examined with a view to making the review process more efficient and more effective. As a general matter, the IRB review process would need to be simplified and unified. At the same time, as noted above, there is a need for enhanced understanding regarding relevant data privacy issues. The consent regime relating to the secondary use of data and biospecimens needs to be re-examined as well. Separately, a trusted third party (TTP) or 'honest broker' could play a crucial role in ensuring data anonymity, while maintaining data quality and integrity at the same time. As such, a further review and examination may be needed as to what, if any, role such a TTP or honest broker could play in the overall data governance regime.

Third, regarding the secondary use of health and medical data, legal and procedural safeguards may need to be re-examined. In particular, in the context of granting access to data or sharing data, appropriate procedural safeguards as well as contractual safeguards need to be in place. Contractual safeguards would include specifying in detail the rights of data subjects and other parties as well as defining the legal relationships among the relevant parties.

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

Acknowledgments

This research was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (No. HI19C1234).

ORCID

Dongjin Lee (http://orcid.org/0000-0002-7816-3737)
Mijeong Park (http://orcid.org/0000-0002-3295-8649)
Seungwon Chang (http://orcid.org/0000-0003-2554-5341)
Haksoo Ko (http://orcid.org/0000-0002-6968-774X)

References

1. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. Health Inf Sci Syst 2014;2:3.
2. Korea Ministry of Government Legislation. Framework Act on Health and Medical Services [Internet]. Sejong, Korea: Ministry of Government Legislation; 2008 [cited at 2019 Oct 1]. Available from: http://www.law.go.kr/lsInfoP.do?lsiSeq=86439&chrClsCd=010203&urlMode=engLsInfoR&viewCls=engLsInfoR#0000.
3. Federal Trade Commission. Protecting consumer privacy in an era of rapid change: recommendations for businesses and policymakers [Internet]. Washington (DC): Federal Trade Commission; 2012 [cited at 2019 Oct 1]. Available from: https://www.ftc.gov/reports/protecting-consumer-privacy-era-rapid-change-recommendations-businesses-policymakers.
4. Organisation for Economic Co-operation and Development. OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data (Paragraph 3: Different degrees of sensitivity) [Internet]. Paris, France: Organisation for Economic Co-operation and Development; c2019 [cited at 2019 Oct 1]. Available from: https://www.oecd.org/internet/ieconomy/oecdguidelinesontheprotectionofprivacyandtransborderflowsofpersonaldata.htm.
5. Council of Europe. Recommendation CM/Rec(2019)2 of the Committee of Ministers to member States on the protection of health-related data [Internet]. Strasbourg, France: Council of Europe; c2019 [cited at 2019 Oct 1]. Available from: https://search.coe.int/cm/pages/result_details.aspx?objectid=090000168093b26e.
6. EUR-Lex. Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the re-use of public sector information [Internet]. Paris, France: EUR-Lex; 2003 [cited at 2019 Oct 1]. Available from: https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=celex:32003L0098.
7. UK National Data Guardian. Caldicott review: information governance in the health and care system [Internet]. London, UK: National Data Guardian; 2013 [cited at 2019 Oct 1]. Available from: https://www.gov.uk/government/publications/the-information-governance-review.
8. UK Biobank. Consent form [Internet]. London, UK: UK Biobank; 2006 [cited at 2019 Oct 1]. Available from: http://www.ukbiobank.ac.uk/wp-content/uploads/2011/06/Consent_form.pdf.
9. Genomics England [Internet]. London, UK: Genomics England; c2019 [cited at 2019 Oct 1]. Available from: https://www.genomicsengland.co.uk/.
10. Legifrance. Law No. 78-17 of 6 January 1978 relating to data, files and freedoms [Internet]. Paris, France: Légifrance; 1978 [cited at 2019 Oct 1]. Available from: https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=LEGITEXT000006068624&dateTexte=19781031.
11. US Department of Health & Human Services. The HIPAA Privacy Rule [Internet]. Washington (DC): US Department of Health & Human Services; c2019 [cited at 2019 Oct 1]. Available from: https://www.hhs.gov/hipaa/for-professionals/privacy/index.html.
12. Hoffman S. Electronic health records and medical big data: law and policy. New York (NY): Cambridge University Press; 2016.
13. Personal Information Protection Commission. Act on the Protection of Personal Information (Act No. 57 of 2003) [Internet]. Tokyo, Japan: Personal Information Protection Commission; 2003 [cited at 2019 Oct 1]. Available from: https://www.ppc.go.jp/files/pdf/290530_personal_law.pdf.

14. Japanese Association of Medical Sciences. Guidelines for genetic tests and diagnoses in medical practice [Internet]. Tokyo, Japan: Japanese Association of Medical Sciences; 2011 [cited at 2019 Oct 1]. Available from: http://jams.med.or.jp/guideline/genetics-diagnosis_e.pdf.
15. European Commission. What is Horizon 2020? [Internet]. Brussels, Belgium: European Commission; c2019 [cited at 2019 Oct 1]. Available from: https://ec.europa.eu/programmes/horizon2020/en/what-horizon-2020.
16. The AEGLE Project [Internet]. Brussels, Belgium: AEGLE; c2019 [cited at 2019 Oct 1]. Available from: http://www.aegle-uhealth.eu/en/.
17. Making NHS data work for everyone [Internet]. York, UK: Healthwatch York; 2018 [cited at 2019 Oct 1]. Available from: https://www.healthwatchyork.co.uk/news/making-nhs-data-work-for-everyone/.
18. NHS Digital. Data sets [Internet]. London, UK: NHS; c2019 [cited at 2019 Oct 1]. Available from: https://digital.nhs.uk/data-and-information/data-collections-and-data-sets/data-sets.
19. About the Portal Epidemiology – France [Internet]. Paris, France: Epidemiology – France; c2019 [cited at 2019 Oct 1]. Available: https://epidemiologie-france.aviesan.fr/en/epidemiology/pages/portal-epidemiology.
20. Commission Nationale de l'Informatique et des Libertés. Un nouveau guide de la sécurité des données personnelles [Internet]. Paris, France: Commission Nationale de l'Informatique et des Libertés; 2018 [cited at 2019 Oct 1]. Available from: https://www.cnil.fr/fr/un-nouveau-guide-de-la-securite-des-donnees-personnelles.
21. US Department of State. Open Government Plan [Internet]. Washington (DC): US Department of State; 2016 [cited at 2019 Oct 1]. Available from: https://www.state.gov/wp-content/uploads/2019/04/Open-Government-Plan.pdf.
22. US Department of Health and Human Services. Guidance regarding methods for de-identification of protected health information in accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule [Internet]. Washington (DC): US Department of Health and Human Services; 2012 [cited at 2019 Oct 1]. Available from: https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html.