Executive Summary (영문)

▢ Background

National Evidence-based healthcare Collaborating Agency (NECA) conducted a study through the collation of prospective/retrospective patient registration data and data with healthcare big data, including claims data and cancer registration data, to generate a basis for comparative effectiveness of various medical technologies and identify the economic value of improved clinical effects of the corresponding medical technology. As a result, a total of 33 types of data were constructed from 2015 to 2017, and clinical research data will be steadily constructed in addition to NECA's existing research projects. In addition, the National Health Clinical Research Project was promoted to establish the basis for national clinical research materials from 2015 to 2017. A total of 17 types of patient registration datasets were collected using iCReaT (web-based Clinical Research and Trial) operated by the National Institute of Health in Korea Center for Disease Control and Prevention. In the future, these clinical research data are expected to be constructed with a large number of various patient data as the national health clinical research project continues. Therefore, it is necessary to diagnose a plan to utilize the established clinical research data efficiently.

▢ Objective

One of the most common problems faced by researchers who perform conventional clinical and epidemiological studies is the lack of a sufficient number of subjects to test their hypotheses or fail to obtain statistically significant results. In this respect, if there is already established clinical research data and it can be used, it is meaningful in terms of utilization cost such as cost reduction and time saving. Currently, the major projects of the Korea Institute for Health and Medical Research are three projects: health care research, new medical technology evaluation, and national health clinical research. In carrying out these research projects, Although the

research material is naturally constructed, there are some considerations to take advantage of it.

First, legal issues related to patients' use of personal medical information

Secondly, is the data that is worth building? The question of the value of the material in terms of content and the fulfillment of the physical factors such as the accuracy of the data.

Third, how would you use available data? Therefore, practical considerations of the value of the legal review and data is assessed for the diagnosis of the availability of the deployed clinical research data, and the practical use cases are presented based on the assessment of the value of the data.

▢ Methods

Based on the consultation of legal experts with regard to ownership of the constructed clinical data, the law review was conducted. Specific legal review contents are as follows.

- Ownership issues with established clinical data
- Ownership of the output of the generated clinical data
- Problems with storage, security, disposal, etc.

Such legal review could provide direction for the ownership of data and results from future tasks performed by the Korea Institute of Health and Medical Service and the government.

Based on this, the related legal contents are reviewed for possible link between clinical research data and data from other agencies, and the drafts of legislation and amendments are prepared and proposed through consultation with other agencies and government officials. In addition, the organization's role of the NECA is reviewed in terms of the possibility of linking data to agencies, and the direction to be pursued is presented.

This study analyzes the barriers to the use of clinical research data and the current status of barriers related to the use of established clinical research data in addition to legal issues. It also determines whether actual

deployed clinical research data can be applied based on advice from experts and stakeholders from all walks of life. Discuss and review expected effects, potential problems, barriers to resolution, and conduct analyses in terms of PESTL (Policy, Economic, Social, Technical, Legal) on the factors that drive the clinical research data deployed.

▢ Results

- **Secondary use of pre-established clinical research data:** It is difficult to make secondary use within the current legal system. However, it is possible to utilize it if the personal information and the individual identifier are discriminated or deleted
- **Secondary use of clinical research data to be established in the future:** It can be used when it is provided with the consents of third party provision and secondary use agreement.

Ⅰ. Laws and Regulations

1. Significance of Personal Information

(1) The Bioethics and Safety Act divides personal information into personal identification information and personal information. "Personal identification information" refers to information that can identify an individual, such as a subject, a name of a donor of an embryo, an oocyte, a sperm or a human body, a resident registration number, and the like. "Personal information" . In addition, "anonymization" refers to the permanent deletion of personally identifiable information, or the replacement of all or part of personally identifiable information with the unique identifier of that person.

(2) Personal information in the Personal Information Protection Act means information about an individual that can be identified by name, resident registration number and video (including information that can be easily combined with other information, even if it is not possible to identify a particular individual).

2. Informed consent for clinical resereach

Prior to conducting clinical studies on human subjects, Article 16 of the Bioethics Act requires written consent (including electronic documents, including the following) to be included in the study, including the following, and does not require consideration of the Personal Information Protection Act.

3. Collecting, using, and providing personal information

(1) In the process of consent for the use of personal information, consideration shall be given to the application of the Privacy Act and the Bioethics Act when the provision of third parties and the secondary use of personal information is intended.

(2) If the purpose of secondary utilization is clear, the information body shall obtain consent from the information subject according to Article 16 of the Bioethics Act and the information provided pursuant to the provision of personal information shall be agreed. In such a case, the collection/ utilization of personal information in Article 15 of the Personal Information Protection Act and the provision of personal information in Article 17 shall be presented separately, and the information shall be notified to the information subject, and if any changes are made, the information shall be notified and agreed to.

(3) Where the purpose of secondary use is not clear, it may be provided on the basis of the review of the institutional committee by article 18 of the Bioethics Act (providing personal information) for secondary research other than the original purpose. However, it is not subject to deliberation at the institutional review board for purposes other than research, i.e. for commercial purposes, so this requires that the explicit purpose of data provision be explained and agreed to by the informant.

4. Privacy

(1) To handle unique identification information under the Personal Information Protection Act, it shall be obliged to secure safety of personal information protection and to protect secrets under section 24.

(2) In particular, the personal information collected shall explain in detail the location of storage, storage period, and disposal method in accordance with Article 21 of the Personal Information Protection Act and Article 16 of the Enforcement Decree of the Personal Information Act.

Ⅱ. Utilization of clinical research data

1. Established clinical research data

There may be methods for anonymizing or deciphering the contents of personal information through expert advice, but there are limitations that information value will decrease if it is not identified. Another option is to push for revision of the content to Article 25 of the Health and Medical Technology Promotion Act, "Medical Health Institutions and Health Institutions under the Act," but there are practical limitations. In addition, the problem with the instrument axis data is that the IRB approval period varies from one study to another, and the IRB approval is valid during that research period. Therefore, secondary utilization requires the approval and management of IRB so that research can be carried out over a long period of time through the IRB of a higher authority or a national institution that acts as a data processing agency or a public IRB. Therefore, for data sources already collected, the possibility of consent exemptions (the reason for exemption of consent under Article 16.3 of the Bioethics Act) can be checked, and individual identifiers can be deleted and non-identifiable from the perspective of 'Future Data'. However, it is also possible to consider ways to collect non-identification as a national resource after establishing an association with other sources or secondary sources (e.g., claims, etc.) because the value of data utilization decreases if individual identifiers are not available.

2. Clinical research data to be established in the future

There are three considerations for the construction and secondary utilization of clinical studies to be developed in the future.

(1) To collect national clinical research data, it is necessary to study sample design and research design to gather participants of the study based on the consent of participants, purpose of research, characteristics of variable names and variables collected in research The protocol should be thoroughly reviewed before it starts.

(2) In addition to collecting data, a committee, such as a publication committee, should be established to review and apply the research topic and to review the data disclosure. It is necessary to operate a committee that can carry out surveillance roles together.

(3) Before data are released, it is necessary to review data cleaning, which is one of the preparatory processes. It is necessary to review the data cleaning process related to the data of the health insurance corporation and health insurance evaluation Based on sufficient discussion and provision of procedures, it is necessary to fully examine the limitations of the possibility of the secondary use of research data collected for primary purposes.

3. Security of data

In order to utilize the data in future, it is necessary to construct analytical room and online platform of the relevant institute. In order to do this, it is necessary to provide guidelines and guidelines for the convenience of users as well as the obligation of data processor to manage them. Also, regular training of established database holders is needed.

▢ Conclusions and Plain Language

First, it should be done firstly to establish governance for secondary utilization. Medical information can be used for public purposes as well as for personal information that should be specially protected. The established clinical research data is sensitive information, and for data intensive research, data processing of sensitive information is necessary and it should

be done properly. Therefore, it is necessary to manage the risk of data by the personal information processor even if it is secondary use. And it is necessary to create an organization that also requires personnel at the privacy risk management level and serves as research ethics and interests management, education, etc.

Next, improvement of IRB for research purposes is necessary. To make the use of medical information more seamless than it is today, it is necessary to prepare different data management requirements for each purpose and process, such as direct care, research and insurance claims, and so on. What is needed at the research site will be documented in each research plan, such as the scope, characteristics, and connections of the data to be utilized at the research planning stage, and an agreement among the data holding agencies will be required for sharing or linking the data. To ensure public interest, the conditions of the research project may be limited, or researchers may sometimes define the qualification of their own agency.

Finally, the contents of medical information utilization and data ownership. One of the ways to address legal issues about data ownership is to consider using clinical research data built as public data. In order to have a status as a public resource, it is necessary to review the scope of data requested by public institutions such as the Korea Health and Medical Research Institute on behalf of researchers, and to establish procedures for consultation with the data holding organization.

▢ Acknowledgement

This Research was supported by National Evidence-based Healthcare Collaborating Agency (NECA) funded by the Ministry of Health and welfare (grant number NP18-004).

Key words: clinical research data, secondary utilization, bioethics law, privacy law, public data, data linkage