Research on Personal Medical Information Protection Based on Big Data

Yi Bu
Shenyang University of Technology, Shenyang, China

Abstract. The rapid development of big data, cloud computing, and artificial intelligence has brought new opportunities to the development of the medical field. The emergence of electronic medical records and medical information systems has provided great convenience for people's medical treatment. However, the arrival of the big data era also poses new challenges to the protection of human medical privacy. The development and application of medical big data have led to explosive growth of data in the medical field, putting pressure on the management of medical institutions. The sharing of medical data between medical institutions and between medical institutions and third parties also poses significant security risks. Moreover, medical data leakage incidents have been continuously exposed in recent years, seriously infringing on the identity and privacy rights of patients. This article aims to explore the research on personal medical information protection based on big data. Firstly, the importance and necessity of protecting personal medical information were analyzed, as well as the current application status and existing problems of big data in medical information management. Secondly, a comparative analysis of domestic and international research on medical privacy protection was elaborated, including the analysis of medical information privacy protection standards in the United States, Europe, and Asia, and the differences and commonalities of international medical information privacy protection standards were compared and analyzed. Then, the theoretical basis of medical information privacy protection was summarized. Then, the application methods and technologies of big data in personal medical information protection were discussed. Finally, relevant countermeasures and suggestions for the protection of personal medical information in big data were discussed, including privacy protection technologies and policy recommendations, as well as ethical issues and norms for the protection of medical information privacy. Through this study, it is hoped that it can provide reference and inspiration for the protection of personal medical information in big data, and provide ideas and suggestions for future development directions.

Keywords: big data; human medical privacy; protection.

1. Introduction

1.1 Research Background

The application of big data in medical information management has achieved certain achievements, including but not limited to clinical decision support systems, disease prediction and prevention, medical resource management and optimization, etc. However, with the continuous development of big data technology, there is also a risk of medical information privacy leakage, which poses new challenges to the protection of personal medical information. Currently, there are issues with the application of big data in medical information management, such as data security and privacy protection, which require further research and improvement.

1.2 Research significance

Although there has been considerable research on the protection of patient privacy rights in our country, due to the fact that medical big data involves more personal privacy and the harm caused to patients is greater, it is necessary to explore the new problems faced by patient privacy protection in the context of social development. From the perspective of patients' privacy, the previous research on patients' privacy is still based on common scenarios, such as discussing the conflict between the privacy of special diseases such as AIDS and the public's right to know, or partially involving the protection of patients' privacy in the use of electronic medical records. However, there is little analysis...
of the issues faced by patient privacy rights in the context of medical big data. From the perspective of data application, there are not many previous literature that have refined research directions to medical data. The literature on medical data also focuses more on whether medical data has property value and the composition of this special type of data rights. Therefore, it is necessary to explore from a new perspective of combining medical big data and patient privacy rights, how to prevent medical data leakage and illegal use, in order to achieve the protection of patient privacy rights.

2. Development Status and Trends of Medical Privacy Protection

2.1 US Medical Privacy Protection Standards

The United States has a comprehensive legal, regulatory, and standard system in the protection of medical information privacy. Among them, the most representative is the Health Insurance Portability and Accountability Act (HIPAA). This bill stipulates the scope of protection for personal medical information and the privacy protection measures that medical institutions and related organizations must comply with. In addition, the United States has established specialized agencies responsible for overseeing the privacy protection of medical information, such as offices under the Department of Health and Human Services (HHS). In addition, the United States is constantly improving relevant laws, regulations, and standards to meet the new needs and challenges of medical information management and privacy protection.

2.2 European Medical Privacy Protection Standards

The European standards for privacy protection of medical information are mainly reflected in the General Data Protection Regulations (GDPR). This regulation stipulates the scope of protection for personal data and the requirements for data processing, which also includes the protection of medical information. Europe attaches great importance to the privacy protection of medical information, not only requiring medical institutions and relevant organizations to strictly comply with data processing regulations, but also stipulating the rights of data subjects and restrictions on data transmission. In addition, Europe has established a dedicated data protection regulatory agency responsible for supervising and managing the privacy protection of medical information.

2.3 Asian Medical Privacy Protection Standards

There are certain differences in the standards for medical information privacy protection in the Asian region. Some countries and regions have formulated specific laws and regulations for the protection of medical information privacy, such as China's Personal Information Protection Law and Japan's Personal Information Protection Law. These laws and regulations establish requirements for the collection, storage, processing, and transmission of medical information, protecting individual privacy rights. However, compared to the United States and Europe, the overall standards for medical information privacy protection in the Asian region still need to be further improved and unified.

2.4 Comparison of Medical Information Privacy Protection Standards in the United States, Europe, and Asia

The protection of medical information privacy is a focal issue of concern in various countries, and the United States, Europe, and Asia have different legislative and regulatory systems in terms of medical information privacy protection standards. The privacy protection of medical information in the United States is mainly centered around the Health Insurance Portability and Accountability Act (HIPAA), which sets standards and privacy requirements for the protection of medical information. Europe, based on the General Data Protection Regulations (GDPR), has put forward stricter requirements for the privacy protection of medical information, including the legality and transparency of data processing, and the protection of the rights of data subjects. Asian countries also have their own legislation and management systems in terms of medical information privacy.
2.5 Trend analysis of international medical information privacy protection standards

With the development of big data technology and the digitization process of medical information, international medical information privacy protection standards are also constantly evolving and improving. The development trend of future international medical information privacy protection standards is mainly reflected in the following aspects: firstly, strengthening the supervision and cooperation of cross-border data flow, especially the establishment and improvement of privacy protection mechanisms for cross-border medical information data flow; The second is to strengthen the legality and transparency of data usage, standardize the data collection, storage, processing, and sharing behavior of medical information, and strengthen the supervision of privacy protection during the use of medical information data; Thirdly, it emphasizes the protection of the rights of data subjects, including the right to access, correct, delete, and oppose personal data, and strengthens the protection of the rights of data subjects in the process of medical information data processing. Therefore, in the future, international medical information privacy protection standards will pay more attention to the legality and transparency of data use, strengthen the protection of the rights of data subjects, and establish and improve privacy protection mechanisms for cross-border medical information data flow. This has important guiding significance for the improvement and development of China's medical information privacy protection standards.

3. Theoretical Basis of Medical Privacy Protection

3.1 Concept of Medical Big Data

Medical big data is a broad concept that has specific classifications including:

3.1.1 Electronic Medical Record Data

This is the data generated during the patient's medical treatment process, including patient basic information, disease complaints, laboratory data, imaging data, diagnostic data, treatment data, etc. These types of data are generally generated and stored in electronic medical records of medical institutions, which is also the main source of medical data. Electronic medical records facilitate the storage and transmission of medical records, but they do not meet the requirements for data analysis. Approximately 80% of medical data is unstructured data composed of free text, which includes not only long textual descriptions but also table fields containing non-uniform text. The transformation of unstructured medical data into a structured form suitable for computer analysis through medical natural language understanding technology is the foundation of medical big data analysis. The data collected in electronic medical records is the largest and most valuable medical data. By integrating with clinical information systems, the content covers various clinical datasets within the hospital. In the interconnection of electronic medical records, due to their respective interests (limiting patient referrals), major electronic medical record companies are also unwilling to make data interconnection. According to relevant reports from the US government, the sharing rate of electronic medical records is only about 30%.

3.1.2 Inspection data

Hospital testing institutions have generated a large amount of diagnostic and testing data for patients, as well as a large number of third-party medical testing centers that are also generating data. Test data is a subcategory in the medical clinical subsystem, but it can directly reflect the disease
development and changes of patients through test data. At present, clinical laboratory equipment has developed rapidly. By collecting laboratory data through the LIS system, it can contribute to the early onset, early diagnosis, and correct diagnosis of diseases.

3.1.3 Image data

With the development of database technology and computer communication technology, digital image transmission and electronic film have emerged. Medical imaging data is generated through imaging equipment and imaging information systems, and hospital imaging departments and third-party independent imaging centers store a large amount of digital imaging data. Medical imaging big data is a large-scale, high growth, multi structure, high-value, and accurate collection of imaging data generated by medical imaging equipment such as DR, CT, MR, and stored in PACS systems. Equivalent to the core category of medical big data, LIS (Laboratory Information System) big data and Electronic Medical Records (EMR) belong to the same category. The amount of medical imaging data is very large, with a fast growth rate and a high degree of standardization. Compared with other clinical data, imaging data has the best standardization, formatting, and uniformity, and its value development is also the earliest.

3.1.4 Cost data

Hospital outpatient expenses, hospitalization expenses, single disease expenses, medical insurance expenses, examination and laboratory income, hygiene material income, diagnosis and treatment expenses, management expense ratio, asset liability ratio, and other economic related data. In addition to the income and expenses of medical services, it also includes cost data of medical services provided by hospitals, including cost data of drugs, equipment, and salaries of health personnel. In the DRGs payment model based on disease diagnosis related groups, detailed cost data accounting is required. By calculating a large sample size, establish disease standard costs, strengthen disease cost accounting and refine cost management.

3.2 The connotation of personal medical privacy protection

Personal medical privacy protection refers to the protection of personal medical records, diagnoses, prescriptions, treatments, and other medical information to prevent unauthorized access, use, or disclosure. The leakage of personal medical information may lead to infringement of personal privacy rights, and may even be used for illegal purposes, causing serious harm to individuals and society. Therefore, the importance of protecting personal medical information is self-evident. Currently, with the deepening development of medical informatization, the protection of personal medical information is becoming increasingly prominent, and there is an urgent need for relevant research and discussion of solutions.

At present, there is no unified definition of patient privacy in the academic community, which can be roughly divided into three categories. The first category is that scholars believe that patient privacy is personal information that involves the patient themselves and does not involve the public interest of others. The second type of scholars mainly define the medical information related to patients as privacy, believing that patient privacy refers to the patient's personal condition, etiology, diagnosis and treatment information, etc. The third type of scholars believe that patient privacy only involves personal information that patients are legally informed by medical personnel or institutions during the process of receiving medical services in medical institutions due to the needs of diagnosis and treatment services, and cannot be illegally disclosed.

4. Main Risk Reasons for Medical Privacy Leakage

4.1 Clinical research data security risks

Clinical research data generally refers to scientific research initiated by hospitals, academic research institutions, and medical enterprises, mainly used for drugs, medical devices, and medical
Advances in Engineering Technology Research CVMARS 2024
ISSN:2790-1688 Volume-11-(2024)
diagnosis, involving basic demographic data, diagnostic information, case and patient reports, and other data information. The doctors, patients and relevant information involved in clinical research are faced with many data security risks when they are transmitted through special lines, Internet lines and other channels, or when they are stored and used in medical institutions.

4.2 Remote medical data security risks
When conducting remote medical activities for patients, communication, computer, and related network technologies are often required, which may involve hospitals, patients, as well as third-party providers of remote diagnosis and treatment equipment and network operators. During this process, personal privacy related health and medical information such as patient test reports, diagnostic results, and medication information will also be involved. If the remote diagnosis and treatment network is accessed by unidentified personnel, or if there are viruses, vulnerabilities, and other issues with related servers and terminals, data will face risks such as theft, tampering, and malicious uploading.

4.3 Data security risks in medical centers
Medical centers involve cross institutional data aggregation, storing a large amount of data information including basic demographic data, medical record data, health record data, etc. If corresponding storage security mechanisms are not established, they will face data security risks such as illegal login, unauthorized access, abnormal retrieval, impersonation queries, batch theft, and plaintext leakage.

5. Methods and Suggestions for Medical Privacy Protection

5.1 Overview of privacy protection technologies for medical big data
Medical big data privacy protection technology refers to the use of technological means to protect the privacy and security of medical big data, preventing unauthorized access and use. At present, common privacy protection technologies for medical big data include data encryption, data desensitization, access control, security authentication, etc. Data encryption refers to the conversion of medical big data into ciphertext, where only authorized users can decrypt and access the data, effectively protecting the privacy of the data. Data desensitization is the process of processing medical big data through deidentification or anonymization, making personal identity information unrecognizable, thereby protecting the privacy of the data. At the same time, access control and security authentication can restrict data access permissions, ensuring that only authorized users can access medical big data, thereby preventing data leakage and abuse. By comprehensively utilizing these technological means, the privacy and security of medical big data can be effectively protected.

5.2 Big data analysis methods for medical information privacy protection
In the protection of medical information privacy, big data analysis methods can discover potential privacy risks and security issues through the analysis and mining of medical big data, and then take corresponding protection measures. Big data analysis methods include data mining, privacy risk assessment, behavior analysis, etc. Data mining refers to the use of algorithms and models to mine potential information and patterns in medical big data, identify potential privacy leakage risks, and provide data support for privacy protection. Privacy risk assessment is a quantitative analysis of the degree of privacy risk in medical big data, providing a basis for the formulation of privacy protection strategies. In addition, behavior analysis can monitor and analyze the access and usage behavior of medical big data, timely detect abnormal behavior, and take corresponding security measures. These big data analysis methods can effectively assist in the protection of medical information privacy, improving the effectiveness and level of privacy protection.
5.3 Privacy protection technology and policy recommendations

Privacy protection technology and policy recommendations are important means to protect the security of personal medical information. Firstly, regarding privacy protection technology for medical big data, it is recommended to use data encryption and desensitization techniques to ensure that medical information is not illegally obtained during transmission and storage. Secondly, a sound medical information privacy protection policy should be established to clearly define the collection, use, and sharing permissions of medical data, ensuring that personal privacy is not violated. In addition, it is necessary to strengthen information security training for medical institutions and related personnel, enhance their awareness of medical information protection, and reduce the possibility of information leakage from the source. Finally, government departments should establish a supervision and inspection mechanism to regulate the secure use of medical information and ensure effective protection of the privacy of medical big data.

5.4 Moral and Ethical Exploration of Medical Information Privacy Protection

The protection of medical information privacy involves ethical and moral issues, and it is necessary to strengthen the ethical and ethical exploration of medical information privacy protection beyond technology and policies. Firstly, medical institutions and relevant personnel should abide by professional ethics and ethical standards when handling medical information, and must not violate the privacy and right to know of patients. Secondly, the use and sharing of personal medical information should follow ethical principles, respect the patient's right to independent choice and information, and not disclose or abuse medical information without authorization. In addition, society should strengthen ethical education on the protection of medical information privacy, raise public awareness of personal privacy protection, and promote the legal, reasonable, and ethical use of medical information. Finally, government departments should establish a medical information ethics review mechanism to conduct ethical evaluations of technologies and policies related to medical information privacy, ensuring that the use of medical information complies with ethical norms and social ethical values.

6. Conclusion and Outlook

6.1 Research Summary

This article aims to explore the research on personal medical information protection based on big data. Firstly, the importance and necessity of protecting personal medical information were analyzed, as well as the current application status and existing problems of big data in medical information management. Secondly, a comparative analysis of domestic and international research on medical privacy protection was elaborated, including the analysis of medical information privacy protection standards in the United States, Europe, and Asia, and the differences and commonalities of international medical information privacy protection standards were compared and analyzed. Then, the theoretical basis of medical information privacy protection was summarized. Then, the application methods and technologies of big data in personal medical information protection were discussed. Finally, relevant countermeasures and suggestions for the protection of personal medical information in big data were discussed, including privacy protection technologies and policy recommendations, as well as ethical issues and norms for the protection of medical information privacy. Through this study, it is hoped that it can provide reference and inspiration for the protection of personal medical information in big data, and provide ideas and suggestions for future development directions.

6.2 Development Trends and Prospects

In the future, with the continuous development and application of big data technology, personal medical information protection will face more challenges and opportunities. In terms of technology, privacy protection technology will become more intelligent and personalized to meet the constantly
changing needs of medical information management and privacy protection issues. At the same time, policies and regulations also need to be continuously improved and innovated, and relevant laws and regulations need to be formulated in a timely manner to ensure the privacy and security of personal medical information. In addition, moral and ethical issues related to the protection of medical information privacy will also become an important direction for future research, requiring more scholars and experts to invest in exploring and regulating. Taking into account the above factors, the future development of big data personal medical information protection will become more diversified and complex, requiring interdisciplinary and interdisciplinary cooperation and research to promote the development and progress of related fields.

References