Research on the Design of Research Data Platform based on Big Data of Psychiatry

Miao Gao*
Xi’an Medical University, China, 710000

*Corresponding author e-mail: lxh12345678900@outlook.com

Abstract. Big data of health care is an important basic strategic resource of the country. The application of big data of medicine is of great significance for clinical medicine research, scientific management and the transformation and development of medical service mode. Based on the common data model (common data model), the data is cleaned, stored and reconstructed, and the clinical data center system is established for clinical and scientific research analysis. Mainly from the system architecture design, key technology, implementation steps to introduce the system design and implementation, and from the safety and reliability of the system design.

Keywords: psychiatry, big data, research platform

1. Introduction
One of the national strategic guidelines is big data construction, in which the national health big data is an important part of big data construction. The document of the general office of the State Council, guiding opinions on promoting and standardizing the application and development of health care big data, clearly points out that health big data is a basic strategic resource. The collection and application of health big data is of great significance as the transformation of clinical medicine, health management and medical services in the next two decades [1]. Especially for psychiatry, it is more suitable to use clinical big data for analysis and research, and achieve significant scientific research results.

In recent years, the global development of big data is like tea, which has involved all walks of life. With the popularization and application of various information systems in the medical industry. The medical industry precipitated a large number of medical clinical data, with certain characteristics of big data. From the perspective of the development of the whole medical industry, the transformation and application of medical big data will become the trend of the industry development. In 2008, big data: Science in he petabyte era was published in science. Three basic conclusions in IDC Research Report "extracting value from chaos" in 2011 constitute the theoretical basis of big data [2]. The American Public Health Association launched FluNearYou, a website that monitors the spread of influenza and uses big data processing technology to analyze and predict any possible future influenza
outbreaks. However, the research on big data mining of psychiatry is not perfect in developed countries.

At present, the research and application of health big data in China is relatively late compared with other big data research, and there is no scale effect for the time being, but there has been considerable development in the past two years. It mainly drives the development of big data analysis through medical application requirements [3]. For example, Beijing health and Family Planning Commission has begun to analyze epidemic disease outbreaks, public emergencies, population movements and other fields through big data mining and analysis technology, and provide data basis for management through big data analysis. Central South University has also carried out a combination of big data, mobile medical treatment and Internet of things, using Internet of things terminals to collect physiological indicators of human body, transmitting them through mobile network, and finally storing health big data to XiangYa big data center. Similarly, the mining and research of big data for psychiatry is still in its infancy in China [4].

2. System Requirements Analysis

According to the analysis of clinical and scientific research needs and the construction experience of our hospital in clinical research data center and related applications, the contents of this research and development are determined as the following aspects: establishing clinical data center system: collecting and integrating all data of the hospital's clinical information system, including structured data and text format pathological reports, scanning copies of previous medical records, etc., so as to be universal. The common data model is the core, which cleans, stores and reconstructs the data, and establishes the big data center of psychiatric clinical medicine in Guangji hospital for clinical and scientific research analysis.

2.1. Patient index service:
Establish a unique identity mark in the big data platform. It is associated with the information of the original system, such as the visit card, the diagnosis and treatment number, the inpatient number and so on. The identity of the big data system is unique to ensure the availability and effectiveness of the big data.

2.2. Unified user management service:
Associated with other software subsystems and accessed through a unified platform, users can access various medical systems with a unique account password. Easy authorization and user management. On the other hand, the data range can be defined according to the roles and responsibilities of different users and tasks in scientific research projects.
3. System Design and Implementation

![System Architecture Diagram]

**Figure 1.** Schematic diagram of system architecture

The whole psychiatric data center system will adopt the centralized data management architecture to manage the electronic medical record data of each department, including the data extraction of each system in the hospital, the transformation format and classification integration of the extracted data, and the classification storage of the integrated data. The architecture of data processing platform is mainly divided into data layer, component layer, function layer and service layer. The service layer provides support services for the data processing platform, including data extraction, data desensitization, data cleaning and other functions; the function layer is the core part of the platform, mainly managing the core metadata, exchange metadata, log management, node management and other functions; the component layer is mainly used to support the basic parts of the function layer, including database adapter, file adapter, data Data transfer conversion component.

3.1. Medical Ontology Module Design:

Medical terminology dictionary is a standardized and computerized medical terminology to reduce ambiguity. Medical ontology knowledge base is a further standardized medical professional vocabulary. Through the standardization of vocabulary, it provides a reference for the standardization of complex, heterogeneous and unstructured data in the future clinical big data of psychiatry. It enhances the usability and standardization of data, and lays a foundation for the follow-up clinical scientific analysis. In the process of implementation, the lexicon of psychiatry terms will be automatically converted into standard terms of ontology knowledge base through natural language processing technology, and local unique vocabulary concepts will be added to ontology knowledge base to form a more complete vocabulary system. Researchers can input key words to search ontology knowledge base, classify similar vocabulary concepts, form concept set, and then use concept set to help define disease queue.

3.2. Design of Natural Language Processing Module for Text Data:

The case data of the original his system and EMR system are mainly stored in the form of text, while the clinical description in the form of text usually has ambiguity and non-standard description, so it is
an important step in the data processing work to convert these unstructured data into standardized and structured data. Based on the characteristics of psychiatry, we developed the following natural language processing modules to realize data structure: (1) medical named entity recognition module; (2) medical named entity automatic coding module.

3.3. Safety and Reliability Design

**Security**: adopt the advanced data information desensitization technology in the world to ensure that the patient's privacy information is not leaked and the clinical trial data is not illegally used, modified or forged. At the same time, ensure the consistency of system data processing and the data is not lost or damaged due to unexpected circumstances, and provide a variety of security inspection and audit means. In terms of system security, firstly, the physical architecture is equipped with security devices such as network gates and firewalls, and the internal and external networks of the hospital are isolated. Real name authentication for network users. At the network level, unlicensed general-purpose ports with potential security risks shall be closed. In the aspect of software, the anti-virus software and strategy to prevent the virus in the near future. Therefore, security settings and precautions are carried out from various aspects to prevent illegal users from intruding into the system and ensure the security and integrity of data. **Reliability**: the platform has also made multiple preparations in terms of the reliability of use. First, it adopts a variety of high reliability software architecture and high availability technology from the technical level to ensure the normal operation and stability of the system. Especially for the uninterrupted operation of key business and the operation guarantee of abnormal conditions. Backup the database of clinical data center to ensure uninterrupted operation and rapid recovery of sudden failure. Ensure the reliability of the system.

4. Summary

Through the big data mining and analysis technology of psychiatry, in the field of big data and precision medical treatment of mental diseases, aiming at various mental diseases, by means of biomedical big data and analysis and statistics of multiple life groups, a scientific research data platform based on big data of psychiatry is constructed to provide a unified and complete number for all kinds of scientific research information applications of hospitals According to the view, it can help to improve the quality of medical services, reduce medical errors, improve the level of clinical research and reduce medical costs.

References

[1] Tai Andy M Y,Albuquerque Alcides,Carmona Nicole E,Subramanieapillai Mehala,Cha Danielle S,Sheko Margarita,Lee Yena,Mansur Rodrigo,McIntyre Roger S. Machine learning and big data: Implications for disease modeling and therapeutic discovery in psychiatry.[J]. Artificial intelligence in medicine,2019,99.

[2] Rutledge Robb B,Chekroun Adam M,Huys Quentin Jm. Machine learning and big data in psychiatry: toward clinical applications.[J]. Current opinion in neurobiology,2019,55.

[3] Response heterogeneity: Challenges for personalised medicine and big data approaches in psychiatry and chronic pain[J]. F1000Research,2018,7.

[4] Na Kyoung-Sae,Han Changsu,Kim Yong-Ku. Big Data and Discovery Sciences in Psychiatry.[J]. Advances in experimental medicine and biology,2019,1192.