© IJCIRAS | ISSN (O) - 2581-5334
May 2024 | Vol. 6 Issue. 12

THE APPLICATION OF BIG DATA IN THE MEDICAL FIELD

Dr. Katherine Ning LI

Assistant Professor,Institute of Sports Science, Xi'an Physical Education University, China

Abstract

With the accumulation of medical data and the development of information technology, big data has brought tremendous changes to the medical field. Medical data has the characteristics of big data, such as large volume, variety, rapid changes, and low-value density. Big data has widespread applications in clinical decision support, healthcare management, public health, drug development, and healthcare informatization. In clinical decision-making, it can predict disease risks, develop personalized treatment plans, and evaluate medical quality. In healthcare management, it helps optimize resource allocation and improve operational efficiency. In public health, it can assess health status, study epidemiology, and formulate intervention measures. In healthcare informatization, it promotes electronic medical records, telemedicine, and the integration and sharing of medical data. However, big data in healthcare also faces challenges such as data silos, privacy protection, data quality, compliance risks, and computing storage. The future will develop towards intelligent diagnosis and treatment, regional healthcare interconnection, and digital therapeutics.

Keyword: Medical data analysis, Clinical decision support,Data security, intelligent healthcare,Health mana- gement

1.INTRODUCTION

In recent years, big data technology has brought unprecedented opportunities and challenges to the medical field. The amount of medical data is growing explosively, coming from various sources such as electronic medical records, medical imaging, genomic sequencing, and remote monitoring devices. These massive heterogeneous data reflect the four primary characteristics of big data:[1] large volume, variety, high velocity, and low-value density. By profoundly analyzing and modeling these complex data, valuable knowledge can be discovered to support clinical diagnosis and treatment decisions, healthcare management, public health, and drug development.

2. PRACTICES OF BIG DATA IN HEALTHCARE

2.1Applications in Clinical Decision Support and Healthcare Management

1) Disease Risk Prediction

Medical big data contains many high-dimensional heterogeneous data about patients' demographic information, physiological and biochemical indicators, medical history, lifestyle, genetic information, etc. Through machine learning and artificial intelligence technologies, in-depth analysis and modeling of these data can accurately predict the future risks of patients developing various diseases. This disease risk prediction based on big data can provide a basis for clinical doctors to formulate early intervention measures, achieving precise disease prevention.

Take diabetes as an example. By associating and analyzing a large amount of patient demographic data, biochemical test data, medical history data, medication data, lifestyle data, etc., high-risk factors affecting the onset of diabetes can be discovered, and machine learning algorithms can be used to construct disease risk prediction models. Based on the model, Doctors can analyze whether an individual belongs to a high-risk group and take timely lifestyle intervention measures. Risk prediction models can also be applied to screen high-risk populations, providing ample medical data support for clinical trials and medical research.[2] In some developed countries, disease risk prediction has been widely used in the early prevention of major diseases such as cardiovascular diseases, cancer, and chronic obstructive pulmonary disease.[3-4]

Due to individual differences and other factors, a single disease risk prediction is often challenging to achieve high accuracy. Therefore, in the future, it is necessary to integrate more dimensions of data, construct multimodal risk prediction models, and combine them with rule engines, expert systems, and other technologies to improve the reliability and interpretability of disease risk assessment. At the same time, with the promotion of the concept of precision medicine, personalized disease risk prediction will also become an important development direction.

2) Personalized Treatment Plans
In personalized treatment, the patient's comprehensive information is first collected through multimodal data such as medical imaging, genomic sequencing, and clinical pathology. Then, based on machine learning algorithms, disease molecular subtype identification models, treatment efficacy prediction models, etc., are constructed to tailor the optimal treatment plan for each patient. Taking breast cancer as an example, by comprehensively analyzing the patient's genomic data, pathological data, imaging data, etc.,[5]breast cancer can be subdivided into multiple clinical subtypes, and each subtype responds differently to different chemotherapy drugs, allowing personalized, targeted therapy plans to be recommended for patients.

In addition to disease diagnosis and treatment, drug dosage adjustment must be individualized. Due to differences in factors such as age, genotype, and liver and kidney function, different populations have significant differences in the metabolism of the same drug. Using artificial intelligence models to analyze patients' genetic data, physiological and biochemical data, medical history, etc., can predict personalized drug exposure levels for patients and guide drug dosage optimization, effectively reducing adverse reactions.[6] Currently, personalized drug dosage adjustment models based on big data have begun to be applied in critical areas, such as cancer treatment and organ transplantation.

3) Medical Quality Assessment
By associating and analyzing a large amount of medical service data (such as medical record data, medication order data, nursing records, etc.) and patient outcome data (such as outcome status, complication rates, readmission rates, etc.), key factors affecting medical quality can be discovered. Machine learning and other techniques can be used to construct medical quality assessment models.[8] For example, in surgical risk assessment, by analyzing historical surgical data, factors affecting surgical risk, such as patient factors (e.g., age, disease type) and provider factors (e.g., surgical skill, surgical duration), can be quantified to achieve precise surgical risk assessment, providing a basis for clinical decision-making.

Furthermore, medical quality assessment can be further broken down into different levels, such as hospitals, departments, and medical staff. By analyzing the medical data of different institutions or personnel, their differences in medical quality can be compared, and specific directions for quality improvement can be identified. Currently, some medical quality assessment models have been applied to evaluate hospitals' surgical quality and patient care quality, providing support for performance evaluation and refined quality management.[7]

2.2 Applications in Public Health

1) Health Monitoring and Disease Prevention
Assessing the population's health status is a prerequisite for formulating public health policies and optimizing the allocation of healthcare resources. Traditional health assessment methods mainly rely on population censuses and sample surveys, which are limited by time, scope, and cost. However, medical big data provides new technical means for comprehensive and dynamic assessment of population health status.

On the one hand, by associating and analyzing demographic data, electronic medical record data, examination and test data, etc., key health indicators such as incidence rates, prevalence rates, and mortality rates can be quantified for the population. Taking the elderly population as an example, the prevalence of common geriatric diseases can be analyzed based on electronic medical record data, and combined with resident health record data, medical insurance expense data, etc.,[8] the overall health status of the elderly population can be evaluated. Additionally, correlation models can be established between health status and influencing factors (such as lifestyle, environmental exposure, etc.) to identify existing health risk factors.[9]

On the other hand, the rise of wearable devices and mobile internet has facilitated the active collection of physiological health data from the population. Analyzing the continuously accumulated behavioral data, physiological parameters, etc., the population's physical condition, life habits, stress, etc., are continuously monitored. Taking the elderly population as an example, wearable devices can be deployed to monitor their daily walking steps, sleep duration, heart rate, blood pressure, and other vital physiological indicators, and combined with lifestyle log data; individualized health assessment models can be established. This continuous, dynamic health assessment will help promptly detect potential health hazards.[10-12]

Based on the assessment of population health status, the health development levels of different regions and populations can be horizontally compared. Through comparative analysis, health deficiencies in specific regions or populations can be identified, and potential influencing factors such as socioeconomic and environmental factors can be traced, providing a basis for formulating subsequent interventions. Furthermore, with the deepening of health management, spatio-temporal dynamic tracking and local anomaly analysis will become new development directions for health status assessment.

2) Health Intervention Measures

Epidemiological research and health status assessment support formulating scientific public health intervention measures. By analyzing the health risk data of the population, high-risk groups can be precisely identified and targeted health interventions can be implemented.[13]

Taking obesity as an example, by analyzing demographic data, genetic data, lifestyle habits data, medical examination data, etc., potential risk factors for obesity can be discovered. [14]Based on these data, an obesity risk prediction model can be constructed to screen high-risk individuals and implement personalized health education and exercise interventions. During the implementation of the intervention measures, real-time monitoring of physiological indicators such as body mass index and heart rate of high-risk individuals through wearable devices can help track the intervention effects promptly and make adjustments and optimizations as needed.

2.3 Healthcare Informatization

1) Electronic Medical Record (EMR) Systems

Big data technology is critical in electronic medical record systems, including data storage and management, as well as data analysis and mining.[1] First, big data technology can process massive medical data, including patient medical records, diagnostic reports, imaging data, etc. These data can be effectively stored and managed through distributed storage and processing technologies, ensuring data security and integrity. Simultaneously, big data technology can perform in-depth analysis and mining of the data in electronic medical records, revealing potential patterns and associations. Through data mining techniques, it can assist doctors in detecting early signs of diseases, formulating more effective treatment plans, and predicting patients' health risks.

Electronic medical record systems can also provide personalized medical services and clinical decision support.[7] Based on the results of extensive data analysis, the system can provide personalized medical services for each patient, recommending the best treatment plans and preventive measures based on the patient's medical history, genomic information, lifestyle habits, and other data, thereby improving medical effectiveness. At the same time, big data technology can provide decision support for clinical doctors, assisting them in making more accurate diagnoses and treatment decisions. By analyzing clinical data and research results globally, the system can provide doctors with the latest medical knowledge and guidelines, helping them make more scientific clinical decisions.

2)Telemedicine Systems
Big data technology supports remote doctors in diagnosing and monitoring patients. Through data collected from sensors, medical devices, etc., [15]combined with extensive data analysis, remote monitoring and real-time diagnosis of patients can be realized, allowing timely detection of health issues and prompt action. Moreover, the transmission and diagnosis of medical imaging are crucial in telemedicine systems. Big data technology can help accelerate the transmission and processing of medical images, improving remote doctors' diagnostic efficiency and accuracy, thereby better serving remote patients, especially when medical resources are scarce, or patients cannot visit hospitals.

Big data technology also supports personalized telemedicine services and medical resource allocation. The system can provide personalized medical services for remote patients based on patient's personal health data and medical histories. By analyzing patients' health data, the system can develop targeted health management plans and treatment regimens, enhancing the effectiveness of telemedicine services. Simultaneously, extensive data analysis can help telemedicine systems better manage medical resources like doctors and equipment. According to patients' needs and medical data analysis results, medical resources can be reasonably allocated, improving the coverage and efficiency of telemedicine services. Furthermore, big data technology can assess and monitor the quality of telemedicine services. By analyzing patient feedback data, diagnostic results, etc., the system can evaluate the satisfaction and effectiveness of medical services and promptly adjust and improve service models, thereby enhancing the overall medical standard.

3. CHALLENGES FACED BY BIG DATA IN HEALTHCARE

3.1 Data silos

One of the significant challenges in healthcare is the existence of data silos. Healthcare data is often stored in disparate systems that do not communicate with each other effectively. This fragmentation can hinder the ability to derive meaningful insights from the data and create barriers to data sharing and interoperability. Breaking down these silos requires investment in interoperable systems and standardized data formats to enable seamless data exchange across different platforms and healthcare providers.

3.2 Data security and privacy protection

Healthcare data is susceptible and subject to stringent privacy regulations such as HIPAA in the United States and GDPR in Europe. Protecting patient privacy and ensuring data security are paramount concerns when dealing with big data in healthcare. With healthcare data's increasing volume and complexity, maintaining robust security measures to prevent unauthorized access, data breaches, and cyberattacks is a significant challenge. Healthcare organizations must implement strong encryption, access controls, and audit trails to safeguard patient data while allowing efficient data analysis and sharing.

3.3 Data quality issues

Ensuring the quality of data in healthcare is crucial for accurate analysis and decision-making. However, healthcare data is often plagued by incomplete, inconsistent, or erroneous data entries, leading to inaccurate conclusions and potentially harmful decisions. Addressing data quality issues requires implementing data governance processes, data validation procedures, and data cleansing techniques to improve healthcare data's accuracy, completeness, and reliability. Additionally, incorporating real-time data monitoring and feedback mechanisms can help identify and rectify data quality issues promptly.

3.4 Compliance risks

Healthcare organizations must comply with various regulatory requirements governing patient data collection, storage, and use. Failure to comply with these regulations can result in severe penalties, legal liabilities, and damage to the organization's reputation. Big data analytics in healthcare often involves processing large volumes of sensitive data, which increases the risk of non-compliance with regulations such as HIPAA, GDPR, and FDA guidelines. Healthcare organizations need robust policies, procedures, and technologies to mitigate compliance risks and ensure data handling practices align with regulatory requirements. This includes implementing appropriate consent mechanisms, data anonymization techniques, and audit trails to track data usage and demonstrate compliance with regulatory authorities.

Addressing these challenges requires a concerted effort from healthcare organizations, technology providers, policymakers, and other stakeholders to develop and implement effective strategies for harnessing the full potential of big data while safeguarding patient privacy, ensuring data security, and maintaining compliance with regulatory standards.

4. CONCLUSION

With the accumulation of medical data and the development of information technology, big data has brought tremendous changes to the medical field. It is widely applied in clinical decision support, healthcare management, public health, drug development, and healthcare informatization, enabling disease risk prediction, personalized treatment plan formulation, resource allocation optimization, medical quality assessment, and more. However, big data in healthcare also faces numerous challenges, including data silos, privacy protection, data quality, compliance risks, and computing storage challenges. In the future, we should strive for development towards intelligent diagnosis and treatment, regional healthcare interconnection, and digital therapeutics, further promoting progress and development in the medical field.

REFERENCE

[1] Al Teneiji, Abeer Saleh et al. "Factors impacting the adoption of big data in healthcare: A systematic literature review." International journal of medical informatics, vol. 187 105460. 20 Apr. 2024, doi:10.1016/j.ijmedinf.2024.105460

[2] Woldaregay, Ashenafi Zebene et al. "Data-driven modeling and prediction of blood glucose dynamics: Machine learning applications in type 1 diabetes." Artificial intelligence in medicine vol. 98 (2019): 109-134. doi:10.1016/j.artmed.2019.07.007

[3] Soriano, Joan B et al. "Chronic Obstructive Pulmonary Disease and Big Data." "Sobre la enfermedad pulmonar obstructiva crónica y el big data." Archivos de bronconeumologia vol. 57,2 (2021): 144. doi:10.1016/j.arbres.2020.03.029

[4] Chai H, Lin S, Lin J, He M, Yang Y, OuYang Y, Zhao H. An uncertainty-based interpretable deep learning framework for predicting breast cancer outcome. BMC Bioinformatics. 2024 Feb 29;25(1):88. doi: 10.1186/s12859-024-05716-7. PMID: 38418940; PMCID: PMC10902951.

[5] Low, Kee, et al. "Breast Cancer: The Translation of Big Genomic Data to Cancer Precision Medicine." Cancer Science, vol. 109, no. 3, 2018, pp. 497-506, https://doi.org/10.1111/ cas.13463. Accessed 7 May 2024.

[6] Choi, Heejung et al. "Machine learning models to predict the warfarin discharge dosage using clinical information of inpatients from South Korea." Scientific reports vol. 13,1 22461. 18 Dec. 2023, doi:10.1038/s41598-023-49831-6

[7] Tan, Hung-Jui et al. "Clinical Decision Support for Surgery: A Mixed Methods Study on Design and Implementation Perspectives from Urologists." Urology, S0090-4295(24)00307-8. 30 Apr. 2024, doi:10.1016/j.urology.2024.04.033

[8] Yi, Zhanmiao et al. "Medication adherence and costs of medical care among patients with Parkinson's disease: an observational study using electronic medical records." BMC public health vol. 24,1 1202. 30 Apr. 2024, doi:10.1186/s12889-024-18431-y

[9] Bian, Zilong et al. "Genetic predisposition, modifiable lifestyles, and their joint effects on human lifespan: evidence from multiple cohort studies." BMJ evidence-based medicine, bmjebm-2023-112583. 29 Apr. 2024, doi:10.1136/bmjebm-2023-112583

[10] Brodie, M A et al. "Big data vs accurate data in health research: Large-scale physical activity monitoring, smartphones, wearable devices and risk of unconscious bias." Medical hypotheses vol. 119 (2018): 32-36. doi:10.1016/j.mehy.2018.07.015

[11] Yoon, Minjae et al. "Application and Potential of Artificial Intelligence in Heart Failure: Past, Present, and Future." International journal of heart failure vol. 6,1 11-19. 30 Nov. 2023, doi:10.36628/ijhf.2023.0050

[12] Yuan, Hang et al. "Self-supervised learning for human activity recognition using 700,000 person-days

of wearable data." NPJ digital medicine vol. 7,1 91. 12 Apr. 2024, doi:10.1038/s41746-024-01062-3

[13] Danjieshi Mi, Ningning et al. "Genetic risk, adherence to healthy lifestyle behaviors, and risk of cholelithiasis: A population-based cohort study." Preventive medicine vol. 182 (2024): 107942. doi:10.1016/j.ypmed.2024.107942

[14] Chang, Qing et al. "Healthy Lifestyle and the Risk of Metabolic Dysfunction-Associated Fatty Liver Disease: A Large Prospective Cohort Study." Diabetes & metabolism journal, 10.4093/ dmj.2023.0133. 19 Mar. 2024, doi:10.4093/ dmj.2023.0133.

[15] D'Auria, Daniela et al. "An intelligent telemonitoring application for coronavirus patients: reCOVeryaID." Frontiers in big data vol. 6 1205766. 18 Sep. 2023, doi:10.3389 / fdata. 2023.1205766