

Research on AI Applications in Personalized Medicine

Yiyi Rui

Hefei No.8 Middle school, Anhui, China

Abstract: With the rapid development of artificial intelligence technology, personalized medicine has become an important trend in the field of modern medicine. This paper aims to explore how artificial intelligence can achieve more accurate diagnosis and treatment customization by analyzing individual genetic, environmental and lifestyle data. Through a review of existing research and technological progress, this paper reveals the great potential of AI in promoting the development of personalized medicine, and discusses its challenges and future prospect.

Keywords: Artificial Intelligence Technology; Medicine

1. Introduction

Personalized medicine refers to a new medical service model that tailors diagnosis and treatment plans for each patient based on their unique physiological characteristics, genetic background, and lifestyle habits. This model helps to improve the treatment effect of diseases and reduce side effects. In recent years, the widespread application of artificial intelligence technology in the medical field has made it possible to achieve this goal. This section will introduce the basic concepts of personalized medicine, its development history, and its position in current medical practice.

1.1 Concept and Characteristics of Personalized Medicine

The core of personalized medicine is "personalization", that is, diagnosis and treatment are tailored to the specific situation of each patient. It is in sharp contrast to the traditional "one-size-fits-all" medical approach, which usually uses standardized treatment methods to treat all patients. Personalized medicine emphasizes the formulation of treatment plans based on individual differences of patients (such as genetic variation, biomarker levels, lifestyle habits,

etc.), which can not only improve the efficacy but also reduce the incidence of adverse reactions [1].

1.2 Development History

The concept of personalized medicine was proposed as early as the 1990s, but it has not become a reality until the past decade with the rapid progress of molecular biology, genomics and information technology [2]. In particular, the development of high-throughput sequencing technology has made it easier and more economical to obtain individual genetic information, while big data and artificial intelligence technologies have provided powerful tools for processing these massive data.

1.3 Current Status

Currently, personalized medicine has achieved remarkable results in many fields, including but not limited to cancer treatment, cardiovascular disease management and diagnosis of rare genetic diseases. Many countries and regions have incorporated personalized medicine into their national strategic plans and invested a lot of resources to support related research and clinical practice [2].

2. Application of AI in Personalized Medicine

2.1 Genetic Data Analysis

Genetic information is one of the important factors that determine the differences in individual responses to certain drugs. Using machine learning algorithms, researchers can identify gene variants associated with specific diseases from massive genomic data, thereby predicting individual risk of disease and their response to different drugs. For example, in cancer treatment, AI can guide the selection of the most appropriate targeted therapy by analyzing the gene expression profile of tumor cells [3].

2.1.1 Genomic data mining

Genomic data mining mainly includes three aspects: high-throughput sequencing technology, gene variant detection and genetic risk assessment.

High-throughput sequencing technology provides basic data support for personalized medicine. High-throughput sequencing (HTS), also known as next-generation sequencing (NGS), represents a revolutionary leap in genomic research. It can quickly and parallelly sequence millions or even billions of DNA fragments. This technology is known for its ability to generate large amounts of data at an unprecedented speed and at a relatively low cost compared to traditional Sanger sequencing methods. High-throughput sequencing platforms typically use massively parallel sequencing strategies to perform multiple sequencing reactions simultaneously at multiple spatially independent locations on a single flow cell. The most widely used technologies include Illumina's bridge amplification, Roche 454's pyrosequencing, and Applied Biosystems/SOLiD's ligation sequencing. These advances have not only facilitated comprehensive analysis of the whole genome, but also enabled researchers to explore the transcriptome, epigenome, and metagenome with unprecedented depth and resolution. As a result, high-throughput sequencing has become an indispensable tool in many fields such as medical genetics, evolutionary biology, and ecology, driving significant progress in our understanding of genetic variation, disease mechanisms, and biodiversity.

Gene variant detection is a technique that systematically analyzes genetic variation in an individual's genome. It aims to reveal differences in DNA sequences that may have important effects on an individual's physiological characteristics, disease susceptibility, and drug response. With the development of high-throughput sequencing technology, large-scale genomic data has become possible, providing unprecedented opportunities for in-depth exploration of genetic variation and its biological significance. By using advanced bioinformatics tools and machine learning algorithms, we can extract meaningful information from massive amounts of genomic data, including but not limited to single

nucleotide polymorphisms (SNPs), insertions and deletions (InDels), and structural variations (SVs). These variation patterns not only help reveal the genetic basis of diseases, but also provide key information for the formulation of personalized medicine, drug development, and public health strategies [4]. Through machine learning methods, we can more efficiently identify genetic markers associated with specific phenotypes, thereby promoting our understanding of the functions of genetic variations and ultimately promoting the development of precision medicine. The following is a detailed description of this process:

2.1.1.1. Data acquisition and preprocessing

1) Raw data collection: First, high-quality genomic data needs to be collected, which usually comes from next-generation sequencing platforms (NGS), such as Illumina, PacBio, or Oxford Nanopore Technologies.

2) Quality control: Perform quality assessment on raw sequencing data, remove low-quality reads or bases, and ensure the accuracy of subsequent analysis.

3) Alignment: Use alignment software (such as BWA, Bowtie2, or Minimap2) to align sequencing reads to the reference genome and generate alignment files (such as SAM or BAM format).

2.1.1.2. Variant detection and annotation

1) Variant detection: Use specialized variant detection tools (such as GATK, FreeBayes, or Samtools) to identify SNPs, InDels, and structural variants from aligned reads.

2) Variant annotation: Annotate the detected variants to explain their location in the genome and possible functional effects, such as using tools such as Annovar or SNPEff.

2.1.1.3. Prediction of functional impact

1) Rule-based methods: predict the impact of variants based on known genetic variation databases (such as dbSNP, ClinVar, etc.) and literature.

2) Machine learning-based methods: use trained models to predict whether variants may lead to changes in protein function. These models are usually based on various features, such as conservation scores, changes in amino acid properties, etc.

2.1.1.4. Feature selection and model building

1) Feature selection: select the most representative features from a large amount of genomic data for modeling, such as using

LASSO regression, random forest feature importance scoring, etc.

2) Model training and validation: use supervised learning (such as support vector machines, random forests, deep neural networks, etc.) or unsupervised learning algorithms (such as cluster analysis) to build models, and evaluate model performance through cross-validation and other means.

2.1.1.5. Model application and result interpretation

1) New sample prediction: apply the established model to new sample data to predict the impact of its genetic variation on the phenotype.

2) Result interpretation: interpret the model output, combine clinical information and biological knowledge, and understand the biological mechanism behind genetic variation.

Genetic risk assessment is a scientific method that predicts the likelihood of an individual suffering from a disease by analyzing their genetic information. It uses a combination of molecular genetics, bioinformatics and statistical principles to detect and analyze gene variants, identify genetic markers associated with specific diseases, and assess the individual's risk level of the disease based on these markers. Genetic risk assessment not only helps to detect genetically predisposed diseases at an early stage, but also provides a basis for the design of personalized medical programs, thereby promoting disease prevention, diagnosis and treatment. The application of AI technology for genetic risk assessment is to predict the likelihood of an individual suffering from genetic-related diseases by integrating a large amount of genetic data with clinical information. This method usually involves the use of machine learning algorithms to analyze the variation patterns in genome sequences and their association with specific diseases. By training deep learning models to identify complex genetic risk factors, AI can reveal genetic patterns that are difficult to discover using traditional statistical methods. In addition, AI is able to process large-scale genetic databases, achieve fast and accurate risk prediction, and support dynamic updating of assessment results to reflect emerging genetic evidence. This method greatly improves the accuracy and efficiency of genetic risk assessment,

providing strong support for preventive medicine and personalized treatment strategies.

2.1.2 Pharmacogenomics

Pharmacogenomics mainly includes three aspects: drug metabolism capacity prediction, targeted therapy selection, and drug resistance monitoring.

In the study of drug metabolism capacity prediction, artificial intelligence technology is widely used to analyze the complex relationship between individual genetic characteristics and drug metabolism. By integrating large-scale genomic data and information related to drug metabolism, machine learning algorithms can identify genetic variants associated with the activity of drug metabolizing enzymes. These variants, especially those that affect the activity of key metabolic enzymes such as cytochrome P450 (CYP450), are crucial to understanding the role of individual differences in drug metabolism. Through in-depth analysis of these genetic data, researchers can build predictive models that can not only predict an individual's ability to metabolize specific drugs, but also further guide the personalized selection of clinical medications, reduce the occurrence of adverse drug reactions and improve treatment effects. This genetic feature-based drug metabolism capacity prediction method provides a powerful tool for the development of precision medicine.

Targeted therapy selection refers to the process of selecting the most appropriate drug treatment regimen based on the unique molecular abnormalities of tumor cells. This process aims to minimize the impact on normal cells by accurately attacking the key molecular targets of tumor cells. In recent years, artificial intelligence (AI) technology has played an important role in this field. AI can quickly process large amounts of genomic data, pathological images, and clinical information, helping doctors identify specific gene mutations or biomarkers in patients, and then predict which targeted drugs may be most effective for the patient. Models trained by machine learning algorithms can not only speed up diagnosis, but also provide personalized treatment recommendations, allowing doctors to customize the best targeted treatment plan for each patient, thereby significantly improving the success rate of

treatment and the quality of life of patients. Drug resistance monitoring refers to the process of continuously tracking and evaluating the responsiveness of tumor cells to specific treatment plans, which is crucial to ensuring the effectiveness of treatment and timely adjusting treatment strategies. With the development of AI technology, it is now possible to dynamically monitor tumor cell resistance by analyzing a large amount of individualized data. AI algorithms can integrate information from multiple levels such as genomics, transcriptomics, and proteomics, identify biomarkers related to drug resistance, and predict the response pattern of tumors to different drugs. In addition, AI tools can predict the patient's response to anticancer drugs through RNA sequencing data of single tumor cells, thereby helping clinicians to promptly detect the emergence and development trend of drug resistance during treatment, and adjust treatment plans accordingly to ensure that patients receive the most optimized treatment effect. This AI-based monitoring method not only improves the speed and accuracy of treatment decisions, but also provides strong support for personalized medicine [5].

2.2 Consideration of Lifestyle Factors

In addition to genetic factors, living environment and personal habits also greatly affect the pathogenesis and development of diseases. AI systems can integrate data from multiple sources (such as social media, wearable devices, etc.) to assess an individual's lifestyle and its impact on health. This information is crucial for designing more effective prevention strategies, including the following aspects.

Behavioral pattern recognition: AI technology can analyze an individual's lifestyle data, such as diet records, exercise frequency, sleep quality, etc., to identify behavioral patterns that are beneficial or harmful to health. For example, through data such as heart rate and number of steps collected by smart watches, AI algorithms can identify whether an individual's daily activity level meets the recommended standards. In addition, AI can also detect potential mental health problems, such as symptoms of depression and anxiety, from behavioral data on social media and online platforms. These insights can help

detect health problems early and take preventive measures.

Health risk assessment: AI systems can predict an individual's risk of developing certain diseases by comprehensively analyzing multiple sources of data such as genetic information, environmental factors, and personal lifestyle habits. For example, AI models can assess an individual's likelihood of developing cardiovascular disease based on factors such as family history, age, and gender. In addition, AI can also identify early signs of cancer by analyzing imaging data (such as CT scans and MRI images). This precise risk assessment enables healthcare providers to develop more personalized prevention and screening plans.

Intervention recommendations: Based on the above analysis results, AI can provide patients with tailored lifestyle intervention recommendations to reduce the risk of disease or improve existing health conditions. For example, for individuals at high risk of cardiovascular disease, AI may recommend measures such as adjusting diet and increasing physical exercise; for patients with chronic diseases, it may provide personalized rehabilitation plans and medication guidance. AI can even monitor patients' physiological indicators in real time and automatically adjust intervention plans based on their changes to ensure the best treatment effect.

2.3 Comprehensive Personalized Treatment Plan Formulation

By comprehensively considering the above factors, AI can help doctors develop the most suitable treatment plan for each patient. This includes but is not limited to drug selection, dosage adjustment, and non-drug interventions (such as dietary recommendations and exercise plans). In addition, AI can monitor patient responses in real time and dynamically optimize the plan to ensure maximum treatment effect[6].

2.3.1 Treatment Customization

In personalized medicine, multimodal data fusion refers to the process of integrating information from different sources, including genetic information, lifestyle data, and environmental factors, to build a comprehensive and detailed view of the patient. Advanced AI technologies can be used to process diverse medical data. These

technologies include deep learning models such as convolutional neural networks (CNNs) for image analysis and recurrent neural networks (RNNs) such as long short-term memory networks (LSTMs) for time series data analysis. In addition, natural language processing (NLP) techniques are used to mine text information in electronic health records, while graph neural networks (GNNs) are used to model complex biological networks and gene interactions. Through this fusion approach, we can comprehensively analyze patients' genomic characteristics, such as susceptibility gene variants, combined with their daily behavioral patterns (e.g., dietary habits, physical activity) and living environment (e.g., air quality and light intensity in the place of residence). This cross-domain data integration not only enhances our understanding of disease mechanisms, but also reveals patient-specific risk factors, thereby guiding the formulation of customized treatment strategies [7].

In the field of personalized drug recommendation, AI technology is playing a vital role. By integrating multi-source heterogeneous data such as patients' genomic information, clinical history, lifestyle factors, and real-time physiological indicators, AI algorithms can build accurate patient phenotype models. Based on these models, deep learning techniques such as autoencoders and generative adversarial networks (GANs) are used to discover potential biomarkers and disease subtypes, thereby helping to identify patient-specific therapeutic targets. In addition, reinforcement learning methods provide dynamically optimized drug treatment options for each patient by simulating the effects of different drug combinations and their long-term effects. For example, in cancer treatment, machine learning models are used to predict the effectiveness of specific drugs on tumor cells and take into account drug-drug interactions and potential side effects, thereby recommending the most appropriate drug combination for each patient and significantly improving treatment effects, while reducing the risk of adverse reactions. This personalized drug recommendation system not only promotes the development of precision medicine, but also provides strong technical support for the treatment of complex diseases in the future.

Personalized dose adjustment is an important component of precision medicine, which uses the patient's individualized information to optimize drug dosage to achieve optimal efficacy and minimize side effects. This process usually involves collecting information such as the patient's genomic data, physiological parameters, pathological characteristics, and past treatment records. By analyzing this data, combined with advanced machine learning algorithms, predictive models can be built to guide dose selection. For example, using algorithms such as support vector machines (SVM) or random forests, drug responses at different doses can be predicted based on factors such as the patient's metabolic rate, drug sensitivity and tolerance. In addition, closed-loop control systems such as adaptive control theory are also used to dynamically adjust drug dosage to ensure that it maintains optimal levels within the therapeutic window. In this way, doctors can continuously monitor the patient's response during treatment and adjust drug dosage in a timely manner, making treatment more personalized and effective. Personalized adjustment of dosage not only improves the success rate of treatment, but also reduces the health risks caused by over- or under-medication, paving the way for truly personalized treatment.

2.3.2 Dynamic monitoring and optimization

Dynamic monitoring and optimization in personalized treatment plans achieve real-time tracking and evaluation of patients' health status by integrating AI technology and remote monitoring systems. With the help of physiological data collected by wearable devices and mobile applications, AI algorithms can continuously monitor changes in patients' health status and perform predictive analysis based on these data to guide timely adjustments to treatment plans. For example, using deep learning models can extract critical information from complex biomarker patterns, helping to identify early signs of treatment response or risk of potential complications. At the same time, a feedback mechanism is implemented to provide patients with instant feedback through smart devices to encourage them to participate in self-management and improve compliance. In addition, in terms of treatment effect evaluation, AI technology can integrate data from multiple sources, including

imaging data, laboratory test results, and patients' subjective reports, to accurately quantify treatment progress and effects. This AI-based dynamic monitoring and optimization method not only improves the degree of personalization of treatment, but also enhances the interaction between doctors and patients, promoting more efficient and customized medical services [8].

3. Challenges and Prospects

Although AI has shown great potential in promoting personalized medicine, it still faces many challenges. These include data privacy protection, ethical issues, and insufficient interdisciplinary collaboration. In order to solve these problems, government agencies, scientific research organizations and industries need to work together to establish unified standards and specifications, and at the same time strengthen science education for the public.

3.1 Data Privacy and Security

In terms of data collection and storage: Medical data often come from multiple sources, including electronic health records (EHRs), imaging data, laboratory test results, etc. Keeping medical data secure includes protecting highly sensitive personal health information from unauthorized access, ensuring the integrity and security of data throughout its lifecycle, and complying with ever-changing global privacy regulations. Mitigate these challenges by employing advanced encryption techniques, de-identification, and innovative approaches like federated learning. In addition, with the advancement of privacy computing technology, it is expected to achieve more efficient data sharing and analysis without sacrificing data privacy, thereby promoting the development of personalized medicine and bringing more accurate and effective treatment plans to patients.

In order to ensure the safety of patient privacy, data anonymization has become an essential step. This process involves removing or obscuring direct or indirect identifiers in the data through various AI technical means, so that the data can no longer be directly or indirectly linked to a specific individual. For example, differential privacy technology can be used to add random noise to the data set to

protect individual privacy while ensuring that the statistical characteristics of the data are not affected; or data obfuscation technology can be used to reduce the recognition of sensitive data by distorting it, but still Maintain the overall value of your data. In addition, named entity recognition in natural language processing technology can also be used to shield sensitive information in text, and encryption technologies such as homomorphic encryption can be used to ensure the security of data during transmission and processing. These technical means can not only effectively protect patient privacy, but also ensure the availability of data and the validity of research [9].

3.2 Ethical and Social Issues

The application of AI technology in personalized medicine has brought about many ethical and social issues, which involve fairness considerations, transparency and informed consent, and responsibility attribution. First, fairness is the key to ensuring that all patients have equal access to high-quality medical services, but in practice, bias in AI systems may prevent certain groups from receiving fair treatment recommendations; second, transparency and informed consent are the basis for building trust between doctors and patients, but complex AI algorithms are often difficult for non-professionals to understand, making it difficult for patients to fully understand and agree to how their data is used; finally, when AI systems misdiagnose or have other problems, responsibility attribution becomes a thorny issue, and developers, medical institutions, and patients themselves may face uncertain responsibility division. Therefore, solving these problems requires interdisciplinary collaboration, including the joint efforts of ethicists, technical experts, medical workers, and policymakers to ensure that AI technology can responsibly promote the development of personalized medicine [10].

3.3 Technical and Collaborative Challenges

First, building an effective interdisciplinary team is crucial but also complex, because it requires seamless collaboration between medical experts, data scientists, engineers, and ethicists, each of whom needs to contribute

their unique expertise. Secondly, there are also significant barriers to integrating AI systems into existing healthcare infrastructure, such as ensuring interoperability with electronic health record (EHR) systems, maintaining data security, and achieving robust performance in real clinical settings. Finally, continuous innovation is needed to keep up with the rapid development of AI methodology and genomic research, which requires continued investment in R&D and continuous education and training for medical staff. Overcoming these challenges requires the joint efforts of all stakeholders in the healthcare ecosystem to create an environment that promotes interdisciplinary collaboration and technological progress.

4. Conclusion

Artificial intelligence technology is gradually changing the traditional medical model and making personalized medicine a reality. Through the comprehensive analysis of genetic, environmental and lifestyle data, AI can help doctors provide more accurate and effective treatment plans for each patient. In the future, with the further development and improvement of related technologies, we have reason to believe that personalized medicine will bring a healthier and better life to mankind.

References

- [1]Matthew Hutson. How AI is being used to accelerate clinical trials [J]. *Nature*, 2024, 627(8003):S2-S5
- [2]Wang Wenya, Wang Lili, Huang Huiyao, et al. Research on the current status of precision medicine development in China and Europe [J]. *China Medical Guide*, 2024, 26 (01): 2-43
- [3]Zhang Ying. Governance of artificial intelligence (AI) in the field of precision

International Conference on Social Development and Intelligent Technology (SDIT2024)

- medicine [J]. *China Medical Guide*, 2022, 24 (01): 46-49.
- [4]Kevin B. Johnson, Wei-Qi Wei, Dilhan Weeraratne, et.al. Precision Medicine, AI, and the Future of Personalized Health Care. Citation: *Clin Transl Sci*, 2021, 14: 86-93
- [5]Liu, R. et al. Evaluating eligibility criteria of oncology trials using real-world data and AI. *Nature* 2021,592: 629-633.
- [6]Du Yao. Development of intelligent medical diagnosis and treatment decision support system[J]. *China High-tech*,2023 (24): 60-62
- [7]Liu Shaokun, He Zhonglian, Li Bin, et al. Design and application of electronic medical record automatic generation system based on large model[J/OL]. *Chinese Digital Medicine*, <https://link.cnki.net/urlid/11.5550.R.20240731.1444.002>.
- [8]Shen Y, Heacock L, Elias J, Hentel KD, Reig B, Shih G, Moy L. ChatGPT and Other Large Language Models Are Double-edged Swords. *Radiology*[J]. 2023, 307(2):e230163. doi: 10.1148/radiol.230163. Epub 2023 Jan 26. PMID: 36700838.
- [9] Zhang Zhongyuan, Ma Lin, Meng Xiaohu, et al. Analysis of the development trend of the application of intelligent technology in the medical field in my country [J]. *Journal of Medical Informatics*, 2024, 45 (07): 7-13
- [10] Xu Hui, Wang Yi. Analysis of the ethical governance path of medical artificial intelligence under the "hard law-soft law" paradigm [J]. *Journal of Nanjing Medical University (Social Science Edition)* [J/OL], <https://link.cnki.net/urlid/32.1606.c.20240726.1111.008>.