Journal of Biomedical Engineering and Health

Review Paper

Reinforcement Learning in Optimizing Personalized Treatment Plans: A Review

Andrzej Nowak[1], Stencel Krzysztof[2,*]

[1] Faculty of Computer Science and Management, Poznan University of Technology, 60-965 Poznan, Poland
[2] Faculty of Computer Science and Information Technology, Wrocaw University of Science and Technology, Wrocaw 50-370, Poland
*Corresponding author: s.krzysztof@wpias.edu.pl

Abstract. Personalized medicine represents a crucial direction in modern healthcare. Reinforcement learning, an artificial intelligence technique that learns optimal decision strategies through interaction with its environment, demonstrates significant potential in optimizing personalized treatment plans. This paper provides a comprehensive review of research progress in applying reinforcement learning to personalized treatment plan optimization. First, the background knowledge of personalized treatment and reinforcement learning is introduced, elucidating the significance and value of researching this topic. Subsequently, reinforcement learning modeling for personalized treatment problems is explored in detail. An in-depth analysis of algorithmic advancements and technical approaches in reinforcement learning follows. The application domain analysis section examines personalized cancer treatment and chronic disease management perspectives. Challenges faced by reinforcement learning in personalized treatment are discussed, and future development directions are projected. Finally, the paper summarizes the role and significance of reinforcement learning in optimizing personalized treatment plans. Research indicates that reinforcement learning holds great promise in providing robust support for optimizing personalized treatment strategies, thereby advancing progress in the medical field.

Keywords: *reinforcement learning; personalized treatment plans; algorithmic advancements; application domains*

Received on 11 September 2024, Accepted on 19 November 2025, Published on 12 January 2025

Introduction

With continuous advancements in medical technology, demands for higher quality and personalized healthcare services have grown significantly [1]. Traditional one-size-fits-all treatment models struggle to address individual patient differences, giving rise to personalized medicine [2]. Personalized treatment aims to tailor optimal therapeutic strategies based on patient characteristics, disease status, and treatment responses, thereby enhancing therapeutic efficacy, reducing side effects, and improving quality of life [3]. In recent years, the rapid development of technologies such as big data and artificial intelligence has provided new opportunities and tools for personalized treatment [4]. Reinforcement learning, as a sequential decision-making theory and method, can optimize decision strategies through the interactive learning between agents and their environment, offering unique advantages in solving personalized treatment plan optimization problems [5].

Current research on applying reinforcement learning to optimize personalized treatment plans has achieved some progress. Many researchers have applied reinforcement learning to personalized treatments for chronic diseases such as cancer, cardiovascular diseases, and diabetes, yielding preliminary results [6]. However, this field remains in its developmental stage, facing numerous challenges and issues [7-10]. First, in reinforcement learning modeling, accurately formalizing complex medical decision-making processes into reinforcement learning problems remains a critical challenge [8]. Second, regarding algorithmic application, existing reinforcement learning algorithms often face issues such as low computational efficiency and slow convergence when handling large-scale, high-dimensional medical data [9]. Additionally, the interpretability and safety of reinforcement learning algorithms are significant factors limiting their widespread adoption in clinical practice [10]. Finally, regarding data acquisition and utilization, obtaining medical data is constrained by privacy protection and data security concerns. Moreover, data quality varies significantly, with issues such as missing values and noise, posing challenges for training and optimizing reinforcement learning models [11-12].

This paper aims to provide a comprehensive and systematic review of research on reinforcement learning in optimizing personalized treatment plans. Through an in-depth analysis of the current applications, advantages, and limitations of reinforcement learning in personalized treatment, it provides valuable references and guidance for research in this field. Specific contributions include: (1) Summarizing typical modeling frameworks and methods for reinforcement learning in personalized treatment plan optimization, offering clear pathways and frameworks for subsequent research; (2) Providing detailed descriptions and comparative analyses of reinforcement learning algorithm applications in personalized treatment, including algorithm characteristics, strengths and weaknesses, and applicable scenarios; (3) It explores application cases and effectiveness evaluations of reinforcement learning across multiple medical domains such as oncology, chronic diseases, and mental health, demonstrating its potential and value in treating various conditions; (4) It analyzes challenges faced by reinforcement learning in personalized treatment and proposes future development directions and research priorities, offering valuable suggestions for advancing its application in this field.

Reinforcement Learning Modeling for Personalized Treatment Problems

Formalization of Medical Decision-Making Processes

The core of reinforcement learning lies in formalizing decision problems as Markov Decision Processes (MDPs) [13]. In personalized treatment, the MDP is defined as follows:1) States describe the patient's health status and relevant information, including physiological indicators, medical history, genetic characteristics, current symptoms, etc. [13]; 2) Actions represent available treatment options or interventions, such as drug types, dosages, administration schedules, surgical plans, etc. [14]; 3) State transition probabilities indicate the likelihood of a patient transitioning to the next state given the current state and action. Due to the complexity and uncertainty of human physiological processes, these probabilities are often difficult to obtain precisely and require estimation through extensive clinical data and expert experience [15]; 4) The reward function measures the immediate gain or loss from taking a specific action in a given state. In personalized treatment, reward functions are typically linked to treatment efficacy, patient survival time, quality of life, and other factors [16]; 5) The discount factor weighs the relative importance of immediate versus future rewards, with values ranging from [0, 1]. A larger γ value indicates greater emphasis on long-term future benefits, while a smaller γ value prioritizes immediate short-term gains. Based on these definitions, Table 1 summarizes the states, actions, and rewards in the treatment of common diseases.

Table 1. States, Actions, and Rewards in the Treatment of Common Diseases

Disease Type	State	Action	Reward
Diabetes	Blood glucose level Blood pressure Body weight	Drug dosage Dietary advice Exercise plan	Days of blood glucose in target range Complication rate
Lung cancer	Tumor size Stage Gene expression	Chemotherapy plan Radiotherapy dosage Targeted drug	Survival time Quality of life score Toxic reactions
Depression	Symptom score Psychological evaluation Medication adherence	Drug type Psychotherapy plan	Symptom remission degree

As illustrated in Table 1, the abstraction of medical decision-making into well-defined states, actions, and rewards forms the foundational basis for implementing reinforcement learning in personalized treatment scenarios. This structured representation enables the modeling of diverse clinical conditions—such as diabetes, lung cancer, and depression—by capturing the dynamic interplay between patient health indicators, available interventions, and clinically relevant outcomes. In practice, the accurate definition and extraction of state variables and reward metrics are critical for ensuring that reinforcement learning algorithms can effectively interpret patient trajectories and optimize long-term health benefits. By systematically mapping real-world clinical processes into the Markov Decision Process (MDP) framework, researchers and practitioners can leverage reinforcement learning to navigate the complexities of individualized patient care, laying the groundwork for robust, data-driven treatment optimization. The formal structure of a personalized treatment problem in reinforcement learning is typically represented as a Markov Decision Process (MDP). This can be formally defined as:

$$\text{MDP} = \langle S, A, P, R, \gamma \rangle \quad (1)$$

where S denotes the state space, A is the action space, P represents the state transition probabilities, R is the reward function, and γ is the discount factor. This mathematical formalism provides a systematic foundation for modeling sequential medical decision-making, allowing reinforcement learning algorithms to optimize complex treatment plans in a structured manner.

The process of formalizing medical decision-making as a Markov Decision Process (MDP) can be intuitively represented through a structural diagram. Figure 1 illustrates the key components of the MDP framework as applied to reinforcement learning for medical decision-making, encompassing states, actions, state transitions, rewards, and the discount factor. In this context, each patient's health status is abstracted as a "state," while possible clinical interventions, such as medication or surgery, are modeled as "actions." The transition from one state to another, following an action, is governed by transition probabilities that reflect the uncertainty and complexity of human physiology. The "reward" is generally defined according to clinical outcomes, such as symptom improvement or adverse event reduction. The discount factor, meanwhile, is used to balance the importance of immediate versus future clinical benefits. By encapsulating these elements, the MDP framework provides a systematic and flexible approach for modeling the sequential and dynamic nature of medical decision-making, enabling reinforcement learning algorithms to iteratively seek and optimize patient-specific treatment strategies.

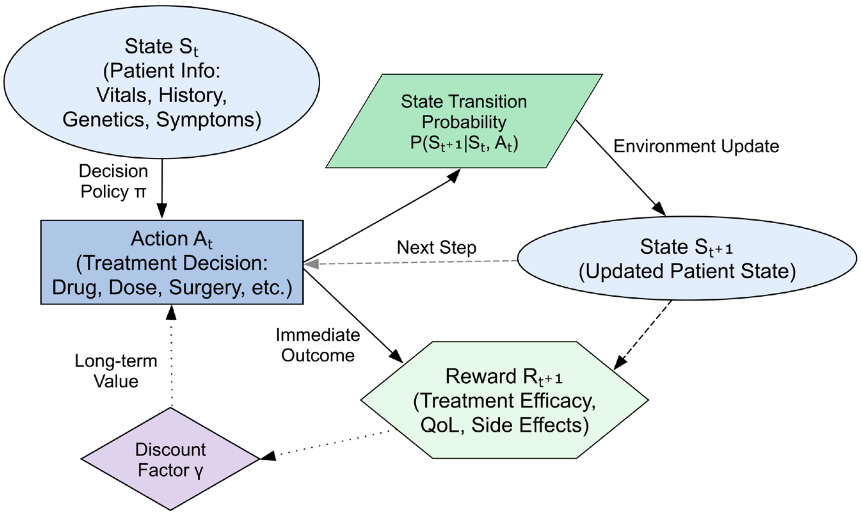

Figure 1. MDP modeling structure for reinforcement learning in medical decision-making.

The structural diagram in Figure 1 visually clarifies the interactions between the various components of the MDP in a clinical setting. By mapping clinical scenarios into the MDP framework, researchers can leverage reinforcement learning algorithms to simulate and analyze complex treatment pathways, account for patient variability, and ultimately optimize medical decisions dynamically over time. This approach not only enhances the interpretability of the decision process but also facilitates the integration of multi-source clinical data, supporting the development of truly personalized and adaptive treatment plans. The diagram serves as a foundational reference for understanding how theoretical models can be translated into practical tools for clinical decision support.

Typical Modeling Frameworks

Common reinforcement learning modeling frameworks for personalized treatment problems include Dynamic Treatment Regimen (DTR), Contextual Multi-Armed Bandit (CMAB), Deep Reinforcement Learning (Deep RL), and Causal Reinforcement Learning (Causal RL). A comparative analysis is presented in Table 2, with specific descriptions as follows:

Table 2. Comparative Analysis of Typical Modeling Frameworks

Modeling Framework	Applicable Scenario	Advantages	Limitations
DTR	Chronic disease treatment requiring sequential decisions	Dynamic adjustment of treatment plans Maximizing long-term treatment effects	Complex model construction Requires large amount of longitudinal data
CMAB	Acute disease treatment with single or few decisions	Simple and fast decision-making Suitable for limited data situations	Cannot consider long-term treatment effects and state changes
Deep RL	Complex disease treatment involving high-dimensional data and complex decision relationships	Strong feature extraction and learning capability Can handle large-scale data	High data and computing resource requirements Poor model interpretability
Causal RL	Scenarios where unobserved confounders affect treatment decisions	More accurate estimation of treatment effects	Affected by empirical reasoning

Each reinforcement learning modeling framework presents unique advantages and is suited to different clinical scenarios in personalized treatment optimization. DTR frameworks excel in managing chronic and progressive diseases by enabling multi-stage, adaptive decision-making, but require abundant longitudinal data for effective learning. CMAB models offer simplicity and efficiency, making them ideal for acute care settings or situations with limited decision points, though they lack the capacity to account for long-term outcomes. Deep RL frameworks, benefiting from powerful feature extraction and representation learning, are highly effective in handling complex, high-dimensional patient data, yet their demand for large datasets and computational resources may limit clinical adoption. Causal RL, by integrating causal inference, addresses confounding and bias issues, thus improving the robustness of treatment effect estimation; however, its success often hinges on the availability of reliable domain knowledge and causal assumptions. Understanding these distinctions is essential for selecting and tailoring the most appropriate modeling approach to specific disease contexts, data environments, and clinical objectives. As research progresses, hybrid frameworks that integrate the strengths of multiple approaches may emerge as promising solutions for the nuanced demands of precision medicine.

To provide a clearer understanding of how these modeling frameworks are applied in personalized treatment, Figure 2 compares the application processes of different reinforcement learning frameworks. This visual overview aids in highlighting the distinctive workflow and decision logic of each approach. By examining the step-by-step progression from patient data collection and state representation to action selection and outcome evaluation, the figure demonstrates how each framework—such as DTR, CMAB, Deep RL, and Causal RL—addresses the unique challenges of clinical decision-making. Moreover, the comparison emphasizes the differences in data requirements, adaptability to dynamic environments, and the ability to handle complex patient heterogeneity. Through this comprehensive illustration, readers can better appreciate how various reinforcement learning paradigms are leveraged to optimize personalized treatment strategies, ultimately supporting more effective, data-driven healthcare solutions.

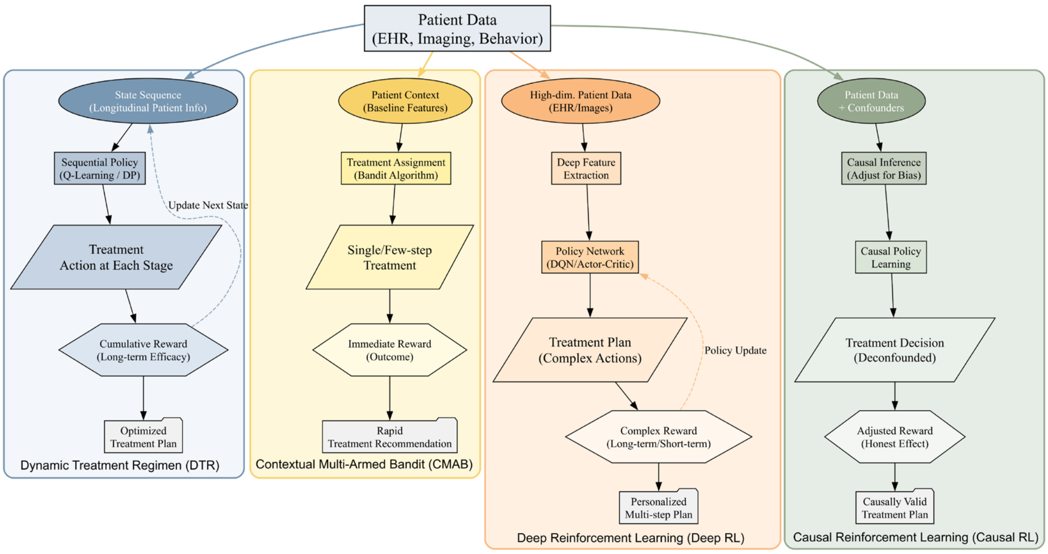

Figure 2. Comparison of application processes for different reinforcement learning frameworks in personalized treatment.

Figure 2 systematically visualizes the distinct mechanisms and optimization pathways inherent to each reinforcement learning framework when applied to personalized medicine. By mapping the flow of information—ranging from patient state acquisition to decision outputs—this diagram underscores the strengths and limitations of each approach in real-world clinical scenarios. Such a comparative perspective not only facilitates the selection of an appropriate modeling framework for specific medical contexts, but also fosters a deeper understanding of how reinforcement learning can be customized to address the complexities of individualized patient care. This graphical summary thus serves as a valuable reference for both researchers and clinicians seeking to integrate artificial intelligence into personalized treatment planning.

Dynamic Treatment Strategy Model

Dynamic treatment strategy is a sequential decision-making framework based on reinforcement learning. It divides the treatment process into multiple stages, formulating corresponding treatment decisions at each stage based on the patient's status information [13]. The core idea of the DTR model is to collect longitudinal patient data and utilize methods such as Q-Learning [17] and dynamic programming [18] to estimate optimal treatment strategies for different stages, thereby maximizing long-term treatment outcomes. Reference [19] applied DTR to address HIV treatment challenges. By dynamically adjusting antiviral regimens based on real-time viral load, CD4 cell counts, and other patient metrics, it improved survival rates and quality of life.

Contextual Multi-Armed Bandit Model

The Contextual Multi-Armed Bandit Model [20] is a simplified reinforcement learning framework suited for single-shot or limited-decision scenarios [21]. In personalized treatment, when decisions involve only a few critical time points or relatively simple regimen choices, the CMAB model can rapidly assign the most effective treatment based on patient context (baseline characteristics, initial condition) [22]. In [23], the CMAB model assigns distinct drug regimens to different patients based on contextual factors like age, gender, and disease severity to enhance treatment success rates.

Deep Reinforcement Learning Models

Deep reinforcement learning combines the powerful representation capabilities of deep learning with the decision-making capabilities of reinforcement learning, enabling it to handle complex high-dimensional state and action spaces [24]. In personalized treatment, deep RL models can learn from large volumes of electronic health records and medical imaging data to automatically extract patient feature representations and formulate personalized treatment plans based on these features [25]. In tumor imaging diagnosis and treatment, deep RL models can analyze patient CT, MRI, and other imaging data to identify tumor characteristics and progression trends, thereby providing precise radiotherapy dosing and treatment plans [26].

Causal Inference Reinforcement Learning (Causal RL) Models

Causal RL is a reinforcement learning model integrated with causal inference methods, designed to address bias issues in traditional reinforcement learning caused by unobserved confounding factors [27]. In healthcare, patient treatment processes are often influenced by multiple unobserved factors such as genetic background and living environment, which may impact treatment outcome evaluation and treatment plan optimization [28].

Data and Simulation Environment

Data plays a crucial role in applying reinforcement learning to optimize personalized treatment plans. Primary sources of medical data include electronic health records [29], medical imaging data [30], wearable device data [31], and clinical trial data [32], with specific source details outlined in Table 3.

Table 3. Research Directions in Medical Imaging Diagnosis

Data Source	Data Features	Advantages
Electronic health records	Rich data dimensions Covering all aspects of patient information	Large sample size Reflects real clinical treatment process
Medical imaging data	Directly shows internal organ structure and function Large information volume	High diagnostic and therapeutic value Provides abundant visual features
Wearable device data	Real-time continuous monitoring of vital signs Strong dynamics	Timely capture of physiological state changes Convenient for long-term monitoring and management
Clinical trial data	Collected under strictly controlled conditions High data quality	High credibility Can be used for validation and model optimization

In addition to these primary data sources, integrating multi-modal medical data has become increasingly important in reinforcement learning for personalized treatment optimization. By combining electronic health records, imaging data, and real-time physiological signals from wearable devices, researchers can obtain a more comprehensive and dynamic representation of patient health status. This holistic data integration enables reinforcement learning models to capture complex relationships between various clinical variables, leading to more accurate and individualized decision-making. Furthermore, the use of high-quality, diverse datasets helps mitigate the risks of data bias and enhances the generalizability of the models across different patient populations and medical conditions. As medical data acquisition technologies continue to advance, the effective utilization and fusion of these data sources will play a crucial role in unlocking the full potential of reinforcement learning in personalized medicine.

Due to numerous restrictions on accessing and utilizing medical data, coupled with inherent risks associated with conducting reinforcement learning experiments in real clinical settings, simulation environments have emerged as crucial tools for researching and testing reinforcement learning algorithms [33]. Common medical simulation environments include physiologically based simulations [34], data-driven simulations [35], and virtual electronic health record systems [36], with comparative analysis presented in Table 4.

Table 4. Comparison of Common Medical Simulation Environments

Simulation Environment Type	Implementation Principle	Advantages
Physiology-based simulation	Built based on physiological principles and mathematical models	Accurately simulates human physiological processes and disease mechanisms Strong interpretability
Data-driven simulation	Mining and modeling based on historical data	Reflects actual clinical data distribution and characteristics Strong adaptability
Virtual EHR system	Generates virtual data by simulating real EHR	Interface and process close to real clinical environment Convenient for research and application

These diverse simulation environments provide essential platforms for developing, evaluating, and validating reinforcement learning models in the context of personalized treatment optimization. By leveraging physiology-based simulations, researchers can gain mechanistic insights into disease processes and therapeutic interventions, facilitating the testing of hypotheses in a controlled virtual setting. Data-driven simulations, on the other hand, allow for realistic scenario modeling based on large-scale clinical datasets, supporting robust algorithm training and validation. Virtual EHR systems further bridge the gap between simulated and real-world clinical applications, enabling reinforcement learning agents to interact with electronic medical record interfaces and workflows similar to those encountered in actual healthcare environments. The integration of these simulation environments not only enhances the safety and feasibility of reinforcement learning research but also accelerates the translation of novel algorithms into clinical practice by providing risk-free opportunities for iterative development and performance benchmarking. As simulation technologies continue to advance, their synergy with reinforcement learning is expected to play a pivotal role in driving innovation and ensuring the reliability of AI-driven personalized medicine.

To further illustrate the integration of various data sources and simulation environments in reinforcement learning for personalized treatment, Figure 3 presents an overview of the data flow and simulation process. This helps clarify how real-world clinical data and virtual environments are utilized in model development and testing. By mapping the pathways from primary sources—such as electronic health records, medical imaging, wearable device data, and clinical trial results—through to data preprocessing, feature extraction, and environment simulation, the figure highlights the complexity and richness of information feeding into reinforcement learning models. It also demonstrates how simulated patient trajectories and outcomes, generated in virtual environments, can be leveraged to supplement real-world data, enhance model robustness, and facilitate safe exploration of novel treatment strategies. Ultimately, this integration supports more reliable and efficient development of reinforcement learning algorithms that are better tailored to the nuances of personalized medicine.

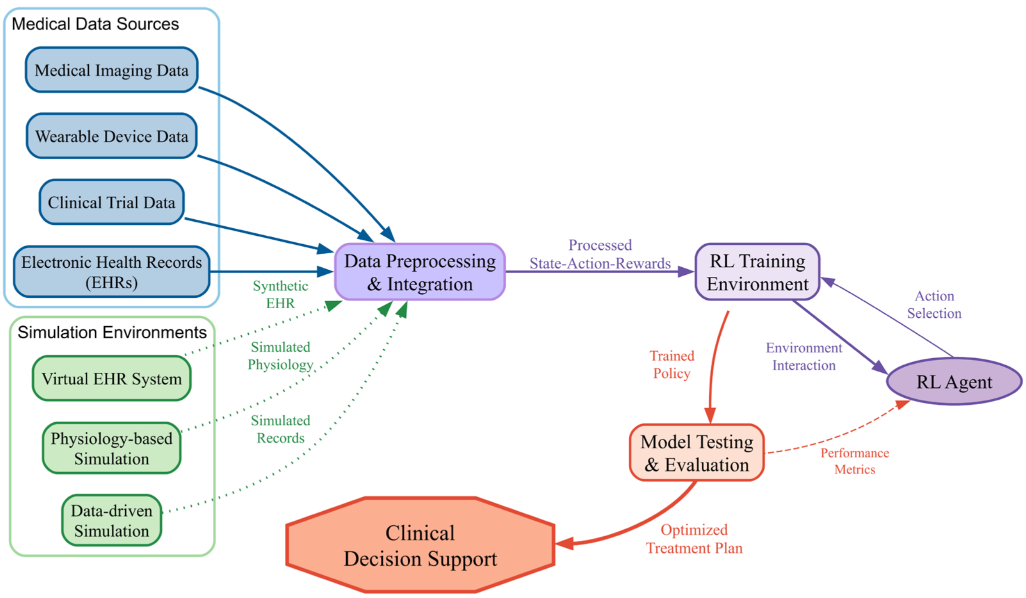

Figure 3. Overview of data flow and simulation environments in reinforcement learning for personalized treatment.

Figure 3 provides a comprehensive visual summary of how diverse data sources and simulation platforms are combined to support reinforcement learning research in personalized healthcare. The diagram underscores the importance of both real-world and synthetic data in overcoming challenges related to data scarcity, privacy, and patient safety. By systematically linking each data type to its respective role in model training, validation, and testing, the figure reveals how a hybrid data-driven and simulation-based approach enables researchers to emulate complex clinical scenarios, evaluate new algorithms, and accelerate the translation of reinforcement learning models into practical, patient-centered treatment solutions. This integrative framework thus paves the way for more robust, scalable, and adaptive AI-driven healthcare innovations.

Reference [34] established a hemodynamic model of the cardiovascular system to simulate the effects of different drug therapies on physiological indicators such as blood pressure and cardiac output, providing a simulation platform for personalized treatment of cardiovascular diseases; Reference [35] constructs a dynamic blood glucose model and drug response model based on extensive electronic medical records of diabetic patients, creating a simulation environment for diabetes treatment. The virtual EMR system mimics real electronic medical record systems, generating virtual medical record data encompassing patient demographics, medical history, diagnoses, and treatments, enabling reinforcement learning algorithms to learn and make decisions under conditions resembling real clinical environments [36].

Algorithmic Advances and Technical Approaches

Value-Based Methods

Value-based reinforcement learning methods determine optimal treatment strategies by evaluating the value function across different state-action combinations [37]. Common approaches include Q-Learning [38] and SARSA [39], with a comparative overview presented in Table 5.

Table 5. Comparison of Value-Based Methods

Algorithm	Update Method	Features
Q-Learning	Greedy policy update	Fast convergence Can find optimal policy
SARSA	Update based on actually executed actions	Stable learning process Suitable for online learning

These value-based methods have demonstrated considerable utility in the healthcare domain, especially when clinical decision processes can be discretized into well-defined states and actions. Q-Learning, with its off-policy nature and ability to converge to the optimal policy independent of the agent's actions, is particularly suited for retrospective analysis of historical patient data, where simulated or real patient trajectories can be leveraged for model training. In contrast, SARSA's on-policy approach makes it advantageous for online and adaptive treatment scenarios, where the agent's policy evolves in real time based on ongoing patient feedback. The choice between Q-Learning and SARSA often depends on the specific clinical application context—Q-Learning may be preferable in environments with abundant historical data and less concern for exploration risk, while SARSA is more appropriate for settings requiring gradual, safe adaptation to new information. As value-based reinforcement learning continues to mature, further integration with deep learning and hybrid methods is anticipated to enhance its applicability to more complex, high-dimensional personalized treatment optimization problems. Value-based methods, such as Q-Learning, update the value of state-action pairs based on observed rewards and predictions for future states. The standard Q-Learning update rule is given by:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha \left[r_{t+1} + \gamma \max_{a'} Q(s_{t+1}, a') - Q(s_t, a_t) \right] \tag{2}$$

Here, $Q(s_t, a_t)$ is the estimated value for taking action a_t in state s_t, α is the learning rate, and γ is the discount factor reflecting the importance of future rewards. This iterative update mechanism enables the agent to learn optimal treatment strategies by continuously refining its value estimates based on patient responses and outcomes.

As shown in Table 6, in personalized treatment, Q-Learning can progressively learn and update the value of different treatment plans based on the patient's current state and historical treatment records, thereby selecting the optimal treatment plan [38]. Reference [40] utilized the Q-Learning algorithm to analyze the patient's symptoms, pulmonary function indicators, and other state information, along with efficacy data under different drug treatment plans, successfully formulating personalized treatment plans for patients and improving treatment outcomes. The SARSA algorithm exhibits greater stability during learning but may converge slightly slower than Q-Learning [39]. In [41], the SARSA algorithm dynamically adjusts medication dosage and treatment plans based on patient status information such as blood pressure readings and medication adherence, effectively controlling blood pressure levels.

Policy Gradient and Actor-Critic

Policy gradient methods directly optimize the policy function by maximizing expected cumulative rewards through gradient ascent [42]. Actor-Critic methods combine the strengths of value-based approaches and policy gradient techniques, simultaneously learning both policy and value functions. These two components collaborate to jointly optimize the decision-making process.

The core idea of policy gradient methods is to parameterize the policy function, representing it as a probabilistic distribution. The policy is then progressively improved by estimating the policy gradient and updating parameters along this gradient [43]. In cancer treatment, treatment selection can be modeled as a policy function that outputs probability distributions for different treatment options based on patient-specific state information, such as tumor characteristics and physical condition [44]. Through policy gradient methods, the parameters of the policy function can be continuously adjusted based on patient treatment responses and efficacy data, enhancing the rationality and effectiveness of treatment plans [45]. Common policy gradient algorithms include the REINFORCE algorithm [46] and the A2C algorithm [47].

In Actor-Critic methods, the Actor selects actions based on the current policy, while the Critic evaluates the quality of the Actor's chosen actions by estimating the corresponding value function [48]. For mental illness treatment, Actor-Critic methods dynamically adjust treatment plans—including medication dosage adjustments and psychotherapy selection—based on symptom changes and psychological assessments to achieve optimal therapeutic outcomes [49]. Representative algorithms include the Deep Deterministic Policy Gradient (DDPG) [50] and Proximal Policy Optimization (PPO) [51], with specific comparisons shown in Table 6.

Table 6. Comparison of Common Policy Gradient Algorithms

Algorithm	Policy Update Method	Value Function Estimation Method	Advantages
DDPG	Gradient-based deterministic policy update	Uses experience replay buffer and target network for value estimation	Handles continuous action space Suitable for complex treatment plan optimization
PPO	Proximal policy optimization Restricts policy update step size	Estimates value function via truncated importance sampling	Stable convergence

The choice between DDPG and PPO for personalized treatment optimization largely depends on the complexity of the clinical scenario and the nature of the action space. DDPG excels in situations requiring fine-grained control over continuous treatment variables, such as precise medication titration or dosage adjustments in chronic disease management. However, it may be more sensitive to hyperparameter tuning and data quality. On the other hand, PPO offers more robust and stable convergence, making it suitable for medical decision processes where safety and reliability are paramount, such as sequential therapy planning or multi-stage interventions. Recent studies have also explored hybrid approaches that combine the advantages of both algorithms, aiming to balance flexibility, efficiency, and stability in real-world clinical applications. As reinforcement learning continues to advance, further research is expected to refine these algorithms for greater interpretability and safety, ensuring their effective and trustworthy adoption in precision medicine.

Figure 4 illustrates the decision-making workflows of Q-Learning and Actor-Critic algorithms in the context of personalized treatment strategy optimization. As described in Section 3.1 and 3.2, Q-Learning, as a value-based reinforcement learning method, updates the value function for each state-action pair through temporal-difference learning, selecting treatment actions based on the learned Q-values. The process involves observing the patient's current health state, selecting a treatment action according to the policy (e.g., epsilon-greedy), applying the intervention, observing the resulting clinical feedback, and updating the Q-table or Q-network to maximize long-term treatment efficacy. In contrast, the Actor-Critic framework combines policy optimization and value estimation, where the Actor generates treatment actions based on the current state, and the Critic evaluates these actions by estimating the corresponding value function. The Critic's feedback continuously guides the Actor's policy update, enabling more efficient and stable learning in complex, high-dimensional medical scenarios. This figure provides an intuitive comparison between the two mainstream reinforcement learning paradigms, highlighting their respective information flows and optimization mechanisms in personalized treatment planning.

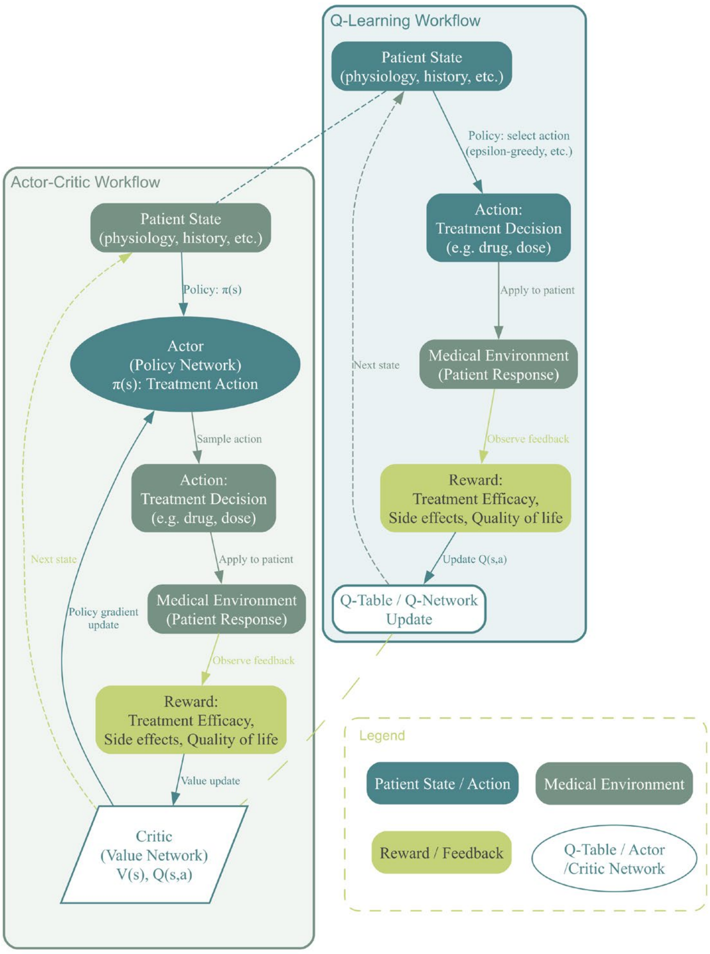

Figure 4. Decision workflows of Q-Learning and Actor-Critic in treatment strategy optimization

Figure 4 offers a visual comparison of the core operational mechanisms underlying Q-Learning and Actor-Critic algorithms within the context of personalized treatment optimization. The diagram highlights how Q-Learning updates value estimates for discrete state-action pairs through an iterative process of observation, action, and reward, making it especially suitable for well-defined and finite clinical decision spaces. In contrast, the Actor-Critic framework separates the policy (Actor) and value estimation (Critic), enabling more flexible learning in complex and continuous action environments. The Critic's feedback guides the Actor's policy adjustments, fostering greater stability and adaptability when optimizing treatment plans for patients with dynamic or high-dimensional health profiles. This side-by-side workflow comparison underscores the strengths and trade-offs of each method and serves as a practical reference for selecting appropriate reinforcement learning strategies to address specific challenges in individualized medical decision-making.

Model-Based RL Algorithms

Model-based reinforcement learning methods simulate patient treatment processes and outcomes by constructing environmental models that learn state transition probabilities and reward functions [52]. Once the environmental model is established, dynamic programming algorithms can be employed to search for optimal treatment strategies [53], or simulated data can be generated to enhance the training effectiveness of model-free reinforcement learning algorithms [54].

In model-based reinforcement learning, the expected cumulative reward-also known as the value function-under a given policy π is defined as:

$$V^{\pi}(s) = \mathbb{E}_{\pi}[\sum_{t=0}^{\infty} \gamma^t r_t \mid s_0 = s] \qquad (3)$$

In this equation, $V^{\pi}(s)$ represents the expected return starting from state s and following policy π, with γ as the discount factor and r_t as the reward at time t. This value function serves as the theoretical basis for assessing the performance of different treatment strategies and guiding the search for optimal personalized interventions in simulated or real clinical environments.

State transition models describe the evolution patterns of patient states under different treatment regimens [55]. When sufficient medical data is available, state transition probabilities can be learned through statistical analysis or machine learning methods [56]. In infectious disease transmission models, information such as patient infection status, exposure history, and treatment measures is used to learn disease propagation patterns within populations and individual patient progression models, providing evidence for epidemic control and personalized treatment [57].

Reward models reflect the therapeutic efficacy and benefits of treatment protocols for patients [58]. Establishing reward models typically requires integrating clinical expertise and experience with patient treatment response data [59]. Reference [60] designed reward functions based on indicators such as rehabilitation training progress and functional improvement, then learned the reward function parameters using machine learning methods. This enables reinforcement learning models to more accurately evaluate the effectiveness of different rehabilitation treatment protocols.

Common planning algorithms include dynamic programming methods such as value iteration [61] and policy iteration [62]. After constructing an environment model, these algorithms efficiently compute optimal treatment strategies. Furthermore, model-based reinforcement learning can be combined with model-free approaches. Environment models generate simulated data to expand training samples for model-free algorithms, enhancing their learning performance and generalization capabilities [63]. In rare disease treatment research, where actual patient data is scarce, disease models can generate simulated data. This simulated data, combined with limited real-world data, can accelerate the discovery of optimal treatment strategies through reinforcement learning model training [64].

Constraints and Safety in RL

Ensuring the safety and efficacy of treatment plans is paramount in personalized medicine. Constraint-based reinforcement learning methods introduce constraints and safety mechanisms to limit the decision space of reinforcement learning models, thereby preventing the generation of treatment plans harmful to patients [65].

Constraints can be defined based on clinical guidelines, expert experience, and patient-specific characteristics [66]. For drug therapy, constraints may include upper/lower limits on drug dosages and contraindications for drug interactions [67]. In surgical treatment, constraints could specify thresholds for surgical risks and preventive measures for postoperative complications [68]. Integrating these constraints into reinforcement learning models ensures generated treatment plans comply with clinical safety standards.

Safety strategy optimization methods incorporate penalty or constraint terms into traditional reinforcement learning objective functions, penalizing behaviors that violate safety constraints. This guides the model to learn treatment strategies that are both effective and safe [69]. Reference [70] employs the Lagrange multiplier method to convert constraints into penalty terms within the reward function, balancing treatment efficacy and safety risks during optimization. Reference [71] employs shielding technology to monitor and correct model outputs in real time. When a recommended treatment plan is detected as unsafe, it is promptly replaced with a safe default plan to ensure treatment safety. A comparative analysis of constrained and safety RL techniques is presented in Table 7.

Table 7. Comparison of Constrained and Safety RL Techniques

Technique	Advantages	Limitations
Constraint optimization	Effectively balances treatment effect and safety risk Stable optimization process	Requires reasonable penalty coefficient setting High requirements for constraint modeling
Shielding	Quickly prevents model from recommending	May reduce model exploration ability Depends on

technique	dangerous actions Ensures treatment safety	expert knowledge for rule design

Incorporating constraint and safety mechanisms into reinforcement learning is especially critical when translating RL-based treatment recommendations into clinical practice. While constraint optimization ensures that generated treatment plans align with established medical guidelines and patient-specific requirements, shielding techniques provide an essential safeguard against the risk of unexpected or unsafe actions during deployment. Together, these methods help build clinician and patient trust in AI-driven decision support systems by prioritizing safety, ethical standards, and regulatory compliance. Moving forward, future research should focus on developing more sophisticated constraint modeling approaches and integrating real-time safety monitoring systems. Additionally, close collaboration with healthcare professionals is necessary to refine constraint definitions and validation protocols, ultimately ensuring that reinforcement learning applications in medicine remain both effective and safe for diverse patient populations.

Meta-Learning and Transfer RL

Meta-learning and transfer reinforcement learning aim to enhance the knowledge transfer capabilities and rapid adaptability of reinforcement learning models across different diseases, patient populations, or treatment scenarios, thereby reducing the costs and time associated with retraining models [72].

The core idea of meta-learning is enabling models to learn how to rapidly acquire new tasks. In personalized treatment, meta-learning can extract universal treatment patterns and prior knowledge by learning from treatment data across multiple related diseases or treatment experiences of multiple patients, forming a “meta-strategy” [73]. Transfer learning facilitates faster convergence and improved performance in the target domain by transferring knowledge and model parameters learned in the source domain. In healthcare, source and target domains may encompass different disease types, patient populations, or medical centers [74]. To better adapt transfer learning to target domain specifics, domain adaptation and personalized adjustments are essential. As demonstrated in [75], transferring treatment plans from adult to pediatric patients requires accounting for pediatric physiology and drug metabolism differences, necessitating model optimization to ensure therapeutic safety and efficacy.

Application Domain Analysis

Personalized Cancer Therapy

Cancer ranks among the major diseases threatening human health, and personalized treatment holds significant importance in oncology. The applications of reinforcement learning in personalized cancer therapy are primarily illustrated in Table 8.

Table 8. Applications of Reinforcement Learning in Personalized Cancer Therapy

Cancer Type	Application Scenario	Reinforcement Learning Method	Main Achievements
Lung cancer	Chemotherapy plan optimization	Deep Q-network	Reduced incidence of adverse effects Improved patient quality of life
Nasopharyngeal carcinoma	Radiotherapy dose and plan optimization	Reinforcement learning algorithm simulation	Improved tumor dose uniformity Reduced damage to surrounding organs
Melanoma	Immunotherapy plan adjustment	Reinforcement learning model analysis	Extended progression-free survival

The adoption of reinforcement learning in personalized cancer therapy marks a significant advancement in precision oncology. RL algorithms can analyze complex, high-dimensional clinical and genomic datasets to tailor treatment strategies for individual patients, optimizing factors such as drug selection, dosage, and scheduling. By continuously learning from patient responses and evolving disease profiles, these models enable dynamic adjustment of therapeutic regimens, potentially improving treatment efficacy and reducing adverse effects. Furthermore, the integration of RL with medical imaging and biomarker monitoring enhances the ability to predict tumor progression and adapt interventions in real time. As research and clinical validation advance, reinforcement learning is expected to become an essential tool in developing adaptive, patient-centered cancer therapies that maximize clinical outcomes and quality of life.

To visually summarize the typical application scenarios of reinforcement learning in personalized treatment across oncology, chronic disease management, and mental health, Figure 5 presents an overview of these domains. This diagram highlights the diversity of use cases and the breadth of clinical applications. It demonstrates how reinforcement learning techniques have been adapted to address the unique challenges within each medical field, from optimizing chemotherapy dosing and radiotherapy planning in cancer care, to improving glycemic control and treatment adherence in chronic diseases such as diabetes and cardiovascular disorders, and further to enhancing rehabilitation outcomes and individualized intervention strategies in mental health and neurological rehabilitation. By mapping these various scenarios, Figure 5 underscores the flexibility of reinforcement learning frameworks in tackling heterogeneous clinical problems, and illustrates the expanding role of artificial intelligence in supporting precision medicine and patient-centered healthcare delivery.

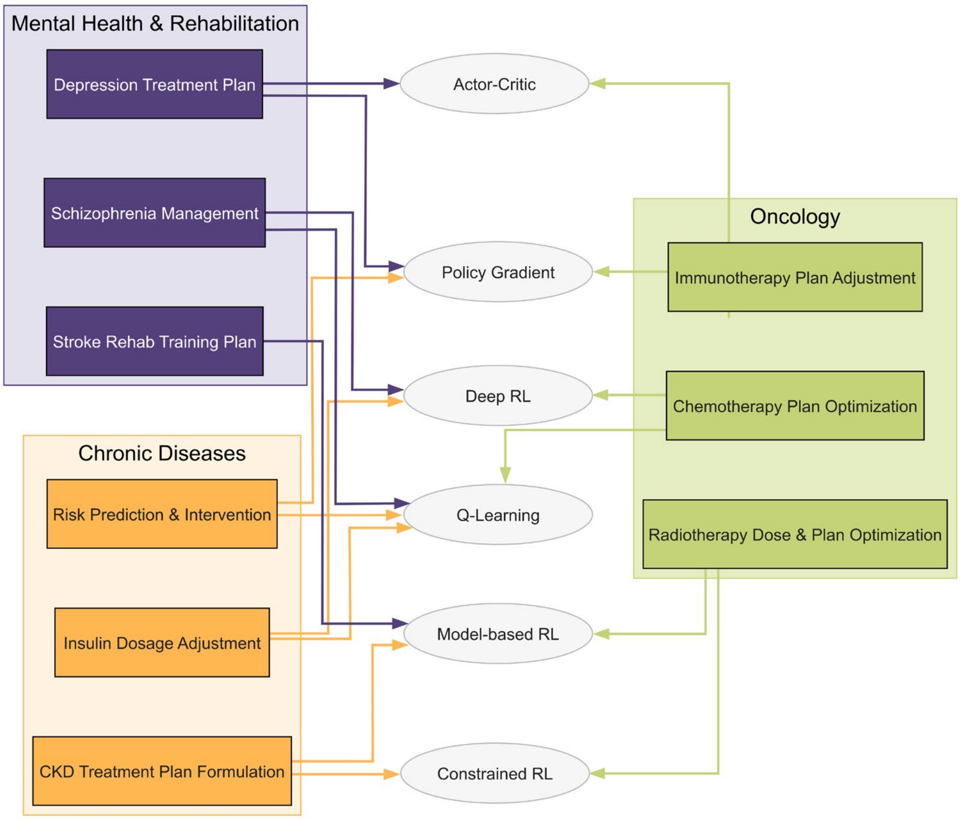

Figure 5. Application scenarios of reinforcement learning in oncology, chronic diseases, and mental health

Figure 5 provides a comprehensive visual synthesis of how reinforcement learning is being implemented across multiple key areas of personalized medicine. By clearly delineating the application domains and corresponding clinical objectives, the figure facilitates an understanding of both the shared and distinct challenges faced in each field. This visual summary not only reinforces the versatility and adaptability of reinforcement learning approaches, but also serves as a valuable reference for researchers and clinicians aiming to identify potential areas for future exploration and cross-disciplinary collaboration. Ultimately, Figure 5 illustrates the transformative potential of reinforcement learning in driving innovation and enhancing the effectiveness of medical decision-making across a wide spectrum of health conditions.

As shown in Table 9, in chemotherapy applications, Reference [76] utilized deep Q-networks to analyze genetic and clinical data from lung cancer patients, optimizing chemotherapy drug dosages and administration cycles. This approach enhanced treatment efficacy while reducing the incidence of adverse side effects in patients. In radiotherapy, Reference [77] employed reinforcement learning algorithms to simulate dose-response relationships between tumors and surrounding tissues under different treatment plans. This enabled personalized radiotherapy planning for nasopharyngeal carcinoma patients, improving radiation dose uniformity while reducing radiation damage to surrounding sensitive organs. In melanoma immunotherapy, Reference [78] employs reinforcement learning models to analyze immune response indicators and tumor shrinkage during immune checkpoint inhibitor treatment. This enables timely adjustments to drug dosage and treatment cycles, extending patients' progression-free survival.

Chronic Disease Management

Chronic conditions such as diabetes, cardiovascular disease, and chronic kidney disease require long-term treatment and management. The application of reinforcement learning in chronic disease management offers new avenues for improving patient quality of life and treatment adherence [79]. Applications of reinforcement learning in chronic disease management are analyzed in Table 9.

Table 9. Applications of Reinforcement Learning in Chronic Disease Management

Disease Type	Application Scenario	Reinforcement Learning Method	Main Achievements
Diabetes	Insulin dosage adjustment	Deep reinforcement learning algorithm	Improved blood glucose compliance rate Reduced blood glucose fluctuation
Cardiovascular disease	Risk prediction and intervention	Reinforcement learning model analysis	Reduced incidence of cardiovascular events
Chronic kidney disease	Treatment plan formulation	Reinforcement learning algorithm learning	Slowed progression of renal function decline

The use of reinforcement learning in chronic disease management offers a transformative approach to long-term patient care. By leveraging continuous streams of patient data from electronic health records, wearable devices, and regular monitoring, RL algorithms can dynamically personalize interventions such as medication dosing, lifestyle recommendations, and risk predictions. This adaptability allows for real-time adjustments in treatment plans, taking into account individual patient responses and changes in health status. As a result, patients benefit from enhanced disease control, reduced complications, and improved quality of life. Going forward, the integration of RL with telemedicine and remote patient monitoring technologies is expected to further expand the possibilities of proactive and individualized chronic disease management, helping healthcare providers deliver more responsive and efficient care.

Mental Health and Rehabilitation

Mental health issues such as depression, anxiety disorders, and schizophrenia severely impact patients' quality of life and social functioning. Traditional treatment methods often struggle to meet patients' individualized needs [80]. The application of reinforcement learning in mental health and rehabilitation offers new approaches to improving patients' mental states and promoting recovery. Applications of reinforcement learning in mental health and rehabilitation are shown in Table 10.

Table 10. Applications of Reinforcement Learning in Mental Health and Rehabilitation

Disease Type	Application Scenario	Reinforcement Learning Method	Main Achievements
Depression	Treatment plan optimization	Reinforcement learning model analysis	Improved treatment response rate and quality of life
Stroke	Rehabilitation training plan optimization	Reinforcement learning algorithm integration	Improved motor function recovery
Schizophrenia	Treatment management plan formulation	Reinforcement learning model analysis	Improved treatment adherence and rehabilitation effect

Recent advances demonstrate that reinforcement learning holds significant promise for addressing the complex and individualized needs of mental health and rehabilitation patients. By adapting to patient-specific feedback and continuously adjusting intervention strategies, RL-based models can support dynamic and adaptive treatment planning, ultimately leading to improved clinical outcomes. Furthermore, the integration of RL with wearable devices and digital therapeutics enables real-time monitoring and personalized support, thereby facilitating earlier interventions and more effective management of mental health conditions. As research progresses, reinforcement learning is expected to play an increasingly pivotal role in developing precision medicine approaches for psychiatric disorders and rehabilitation, offering hope for enhanced recovery and quality of life for affected individuals.

ICU and Emergency Care

ICU and emergency care represent scenarios within the medical field that demand extremely high timeliness and accuracy in decision-making. The application of reinforcement learning in these settings helps improve patient treatment success rates and survival rates [81]. Applications of reinforcement learning in ICU and emergency care are shown in Table 11.

Table 11. Applications of Reinforcement Learning in ICU and Emergency Care

Disease Type	Application Scenario	Reinforcement Learning Method	Main Achievements
Sepsis	Treatment decision support	Reinforcement learning algorithm analysis	Improved patient survival rate Reduced inappropriate antibiotic use
Acute respiratory distress syndrome	Mechanical ventilation parameter optimization	Reinforcement learning model real-time monitoring and adjustment	Improved oxygenation and lung compliance
Cardiac arrest	Emergency plan formulation	Reinforcement learning algorithm simulation	Improved survival rate and neurological recovery

In recent years, with the increasing demand for timely and precise decision-making in critical care and emergency medicine, research on the application of reinforcement learning in ICU and emergency settings has advanced rapidly. Reinforcement learning models can leverage large-scale ICU monitoring data and real-time physiological signals from multiple sources to dynamically optimize treatment parameters and intervention strategies. For example, in sepsis management, reinforcement learning can continuously adjust antibiotic regimens and fluid management plans in real time based on individual patient differences, infection progression, and treatment responses, thereby improving survival rates and reducing the risk of inappropriate antibiotic use. In mechanical ventilation for acute respiratory distress syndrome, reinforcement learning models analyze the relationships between ventilator settings, oxygenation indices, and lung compliance to intelligently adjust ventilation parameters, enhance oxygenation, and reduce the risk of ventilator-induced lung injury. Additionally, for cardiac arrest and other critical emergencies, reinforcement learning can assist clinicians in rapidly formulating individualized emergency plans, improving the success rates of resuscitation and neurological outcomes. Although challenges such as data heterogeneity and insufficient model interpretability remain, the application prospects of reinforcement learning in ICU and emergency medicine are promising, and it is expected to further promote the clinical adoption and development of intelligent decision support systems.

Challenges and Future Directions

Challenges

Although reinforcement learning has achieved some success in optimizing personalized treatment plans, the following challenges remain:1) The decision-making process and rationale of reinforcement learning models are difficult for humans to understand and interpret. In the medical field, physicians and patients demand extremely high standards for the rationality and safety of treatment plans, requiring clear understanding of the basis for treatment decisions and potential risks. 2) Since reinforcement learning models make decisions based on learned data, they may exhibit errors or instability when encountering novel clinical scenarios or shifts in data distribution, potentially compromising patient safety. 3) The application of reinforcement learning in personalized treatment raises numerous ethical and legal concerns, including informed consent, delineation of treatment responsibility, algorithmic bias, and fairness.

Future Directions

To address these challenges, reinforcement learning-based personalized treatment optimization methods should pursue the following development paths: 1) Future efforts should focus on developing more interpretable reinforcement learning models and algorithms. By integrating medical domain prior knowledge and causal relationships through techniques like knowledge graphs and causal inference, models can generate easily understandable decision explanations. 2) Further exploration and development of methods for integrating multimodal medical data, combined with deep learning techniques for deep feature learning, to provide reinforcement learning models with richer and more comprehensive state information, thereby enhancing decision accuracy and personalization. 3) Strengthening the integration of reinforcement learning technology with clinical practice through more clinical trials and real-world application studies to validate the efficacy and safety of reinforcement learning models across different diseases and patient populations.

References

[1] Cilla, S., Romano, C., Morabito, V. E., Macchia, G., Buwenge, M., Dinapoli, N., Indovina, L., Strigari, L., Morganti, A. G., Valentini, V., & Deodato, F. (2021). Personalized Treatment Planning Automation in Prostate Cancer Radiation Oncology: A Comprehensive Dosimetric Study. Frontiers in Oncology, 11. DOI:10.3389/fonc.2021.636529

[2] Lam, C. S., Alagoz, E., Jawara, D., Murtha, J. A., Breuer, C. R., Pati, B., Eierman, L., Farrar-Edwards, D., Voils, C. I., & Funk, L. M. (2023). Patient and Provider Perceptions About Communication After Bariatric Surgery: A Qualitative Analysis. Journal of Surgical Research, 291, 58–66. DOI:10.1016/j.jss.2023.05.029

[3] Chen, Z., Zeng, D. D., Seltzer, R. G. N., & Hamilton, B. D. (2021). Automated generation of personalized shock wave lithotripsy protocols: Treatment planning using deep learning. JMIR Medical Informatics, 9(5), e24721. DOI:10.2196/24721

[4] Harnas, S. J., Knoop, H., Booij, S. H., & Braamse, A. M. J. (2021). Personalizing cognitive behavioral therapy for cancer-related fatigue using ecological momentary assessments followed by automated individual time series analyses: A case report series. Internet Interventions, 25, 100430. DOI:10.1016/j.invent.2021.100430

[5] Mathieu, D., Bedwani, S., Mascolo-Fortin, J., Côté, N., Bernard, A.-A., Roberge, D., Yassa, M., Bahig, H., & Vu, T. (2020). Cardiac Sparing with Personalized Treatment Planning for Early-stage Left Breast Cancer. Cureus, 12(3), e7247. DOI:10.7759/cureus.7247

[6] Al-Medfa, M. K., Al-Ansari, A. M. S., Darwish, A. H., Qreeballa, T. A., & Jahrami, H. (2023). Physicians' attitudes and knowledge toward artificial intelligence in medicine: Benefits and drawbacks. Heliyon, 9(4), e14744. DOI:10.1016/j.heliyon.2023.e14744

[7] Chen, W., Lu, Y., Qiu, L., & Kumar, S. (2021). Designing personalized treatment plans for breast cancer. Information Systems Research, 32(3), 932–949. DOI:10.1287/ISRE.2021.1002

[8] Deluca, C., Beyer, M., Drew, H., Murray, J., Massey, J., Koffman, B., Battiato, K., Gribben, J. G., Lamanna, N., & Russomanno, C. (2023). The Value of Collaboration with Patient Advocacy Organizations to Enhance Initiatives Focused on Patient-Centered Care and Improve Outcomes in Chronic Lymphocytic Leukemia. Blood, 142(Supplement 1), 3715–3715. DOI:10.1182/blood-2023-177959

[9] Arslan, S. (2023). Exploring the Potential of Chat GPT in Personalized Obesity Treatment. Annals of Biomedical Engineering, 51(9), 1887–1888. DOI:10.1007/s10439-023-03227-9

[10] Paap, D., Krops, L. A., Schiphorst Preuper, H. R., Geertzen, J. H. B., Dijkstra, P. U., & Pool, G. (2022). Participants' unspoken thoughts and feelings negatively influence the therapeutic alliance; a qualitative study in a multidisciplinary pain rehabilitation setting. Disability and Rehabilitation, 44(18), 5090–5100. DOI:10.1080/09638288.2021.1924297

[11] Magin, P. (2020). Capsule Commentary on Bolton et al., Integrating Personalized Care Planning into Primary Care: a Multiple-Case Study of Early Adopting Patient-Centered Medical Homes. Journal of General Internal Medicine, 35(2), 621. DOI:10.1007/s11606-019-05577-4

[12] Anchugam, & Jayanthila, D. (2023). Securing healthcare data in smart cities: A machine learning approach. I-Manager's Journal on Information Technology, 12(4), 33. DOI:10.26634/jit.12.4.20368

[13] Li, H., Chen, H., Morgan, L., Li, W., & Oliver, B. G. (2021). A narrative review of clinical studies of herbal treatment of difficult to manage asthma. Complementary Therapies in Clinical Practice, 44(3), 101433. DOI:10.1016/j.ctcp.2021.101433

[14] Ferwerda, B., & Tkalčič, M. (2020). Exploring the Prediction of Personality Traits from Drug Consumption Profiles. UMAP 2020 Adjunct - Adjunct Publication of the 28th ACM Conference on User Modeling, Adaptation and Personalization, 2020-January, 2–5. DOI:10.1145/3386392.3397589

[15] Reiter-Brennan, C., Osei, A. D., Uddin, S. M. I., Orimoloye, O. A., Obisesan, O. H., Mirbolouk, M., Blaha, M. J., & Dzaye, O. (2020). ACC/AHA lipid guidelines: Personalized care to prevent cardiovascular disease. Cleveland Clinic Journal of Medicine, 87(4), 231–239. DOI:10.3949/ccjm.87a.19078

[16] Christian, B., & Karla, A. D. O. (2023). #36482 Revolutionizing nerve pain treatment: harnessing dosimetry, nanobots, and ai for personalized relief. Regional Anesthesia and Pain Medicine, 48(Sup1), A211.1-A211. DOI:10.1136/rapm-2023-esra.386

[17] Chen, H., & Yang, S. (2023). Acute coronary syndrome management in hemophiliacs: How to maintain balance?: A review. Medicine (United States), 102(11), E33298. DOI:10.1097/MD.0000000000033298

[18] Siriwardhana, C., & Kulasekera, K. B. (2020). Personalized treatment plans with multivariate outcomes. Biometrical Journal, 62(8), 1973–1985. DOI:10.1002/bimj.201800072

[19] Mazza, M., Caroppo, E., De Berardis, D., Marano, G., Avallone, C., Kotzalidis, G. D., Janiri, D., Moccia, L., Simonetti, A., Conte, E., Martinotti, G., Janiri, L., & Sani, G. (2021). Psychosis in women: Time for personalized treatment. Journal of Personalized Medicine, 11(12). DOI:10.3390/jpm11121279

[20] Sugandh, F., Chandio, M., Raveena, F., Kumar, L., Karishma, F., Khuwaja, S., Memon, U. A., Bai, K., Kashif, M., Varrassi, G., Khatri, M., & Kumar, S. (2023). Advances in the Management of Diabetes Mellitus: A Focus on Personalized Medicine. Cureus. DOI:10.7759/cureus.43697

[21] Bove, J., Fiala, J. A., Milano, oJrdan, Sandino, A., Asken, B. M., & Bauer, R. M. (2023). 54 Individuals Employing Extreme Coping Behaviors Correlated with Increased Severity of Symptoms Following mTBI. Journal of the International Neuropsychological Society, 29(s1), 159–160. DOI:10.1017/s1355617723002588

[22] Chen, Y., Stewart, J. W., Ge, J., Cheng, B., Chekroud, A., & Hellerstein, D. J. (2023). Personalized symptom clusters that predict depression treatment outcomes: A replication of machine learning methods. Journal of Affective Disorders Reports, 11. DOI:10.1016/j.jadr.2023.100470

[23] Chi, C. L., Wang, J., Ying Yew, P., Lenskaia, T., Loth, M., Mani Pradhan, P., Liang, Y., Kurella, P., Mehta, R., Robinson, J. G., Tonellato, P. J., & Adam, T. J. (2022). Producing personalized statin treatment plans to optimize clinical outcomes using big data and machine learning. Journal of Biomedical Informatics, 128, 104029-. DOI:10.1016/j.jbi.2022.104029

[24] Wen, Z., Wang, S., Yang, D. M., Xie, Y., Chen, M., Bishop, J., & Xiao, G. (2023). Deep learning in digital pathology for personalized treatment plans of cancer patients. Seminars in Diagnostic Pathology, 40(2), 109–119. DOI:10.1053/j.semdp.2023.02.003

[25] Gong, Q., Wang, H., Zhu, M., Wang, H., Ma, T., & Ding, D. (2023). Development of a personalized tinnitus detection and treatment instrument. Qinghua Daxue Xuebao/Journal of Tsinghua University, 63(12), 2076–2084. DOI:10.16511/j.cnki.qhdxxb.2022.25.026

[26] de Geus, E. Q. J., Milders, M. V., van Horn, J. E., Jonker, F. A., Fassaert, T., Hutten, J. C., Kuipers, F., Grimbergen, C., & Noordermeer, S. D. S. (2021). Acquired Brain Injury and Interventions in the Offender Population: A Systematic Review. Frontiers in Psychiatry, 12, 658328. DOI:10.3389/fpsyt.2021.658328

[27] Griffith, L., Donohue, K., Abul-Khoudoud, M., & Gentry, M. (2022). Developing a Pipeline for Personalized Diagnoses of Glut1 Deficiency Syndrome. The FASEB Journal, 36(S1), 3. DOI:10.1096/fasebj.2022.36.s1.r5375

[28] Kollitz, E., Han, H., Kim, C. H., Pinto, M., Schwarz, M., Riboldi, M., Kamp, F., Belka, C., Newhauser, W., Dedes, G., & Parodi, K. (2022). A patient-specific hybrid phantom for calculating radiation dose and equivalent dose to the whole body. Physics in Medicine and Biology, 67(3), 19. DOI:10.1088/1361-6560/ac4738

[29] Xue, W., Sun, C., Yuan, H., Yang, X., Zhang, Q., Liao, Y., & Guo, H. (2022). Establishment and Analysis of an Individualized EMT-Related Gene Signature for the Prognosis of Breast Cancer in Female Patients. Disease Markers, 2022(1), 17. DOI:10.1155/2022/1289445

[30] Pryor, R. R., Pryor, J. L., & McDermott, B. P. (2022). Persistent Knowledge Gaps Regarding Exertional Heat Stroke Treatment. Journal of Athletic Training, 57(8), 756–759. DOI:10.4085/1062-6050-366-19

[31] Salman, A., Metrakos, P., & Bartlett-Esquilant, G. (2022). Tackling topics in primary care: Using consensus-oriented decision-making, a socio-technical method to biobank implementation. Annals of Family Medicine, 20(Sup1), 3089. DOI:10.1370/afm.20.s1.3089

[32] Asadi-Pooya, A. A., & Farazdaghi, M. (2022). Treatment response in newly diagnosed epilepsy: a syndromic approach. Neurological Research, 44(7), 630–635. DOI:10.1080/01616412.2021.2025319

[33] Mohan, V., & Radha, V. (2022). Precision diabetes is becoming a reality in India. Proceedings of the Indian National Science Academy, 88(4), 551–559. DOI:10.1007/s43538-022-00115-z

[34] Guilbert, T. W., Biagini, J. M., Ramsey, R. R., Keidel, K., Curtsinger, K., Kroner, J. W., Durrani, S. R., Stevens, M., Pilipenko, V., Martin, L. J., Kercsmar, C. M., Hommel, K., & Hershey, G. K. K. (2022). Treatment by biomarker-informed endotype vs guideline care in children with difficult-to-treat asthma. Annals of Allergy, Asthma and Immunology, 128(5), 535-543.e6. DOI:10.1016/j.anai.2022.01.030

[35] Cellina, M., Cè, M., Alì, M., Irmici, G., Ibba, S., Caloro, E., Fazzini, D., Oliva, G., & Papa, S. (2023). Digital Twins: The New Frontier for Personalized Medicine? Applied Sciences (Switzerland), 13(13). DOI:10.3390/app13137940

[36] Cerniglia, L. (2023). Personalized Psychological Interventions in Children: Harnessing Advancements in Genetic and Epigenetic Research. Journal of Clinical Medicine, 12(16), 2. DOI:10.3390/jcm12165311
[37] Fiter, R. J., Murphy, L. J., Gong, M. N., & Cleven, K. L. (2023). The impact of air pollution on asthma: clinical outcomes, current epidemiology, and health disparities. Expert Review of Respiratory Medicine, 17(12), 1237–1247. DOI:10.1080/17476348.2024.2307545
[38] Qin, S., Lin, Z., Liu, N., Zheng, Y., Jia, Q., & Huang, X. (2023). Prediction of postoperative reintervention risk for uterine fibroids using clinical-imaging features and T2WI radiomics before high-intensity focused ultrasound ablation. International Journal of Hyperthermia, 40(1), 2226847. DOI:10.1080/02656736.2023.2226847
[39] Yi, H. (2023). Improving cloud storage and privacy security for digital twin based medical records. Journal of Cloud Computing, 12(1). DOI:10.1186/s13677-023-00523-6
[40] Jiang, J., Deng, H., Liao, H., Fang, X., Zhan, X., Wei, W., Wu, S., & Xue, Y. (2023). An Artificial Intelligence-Enabled ECG Algorithm for Predicting the Risk of Recurrence in Patients with Paroxysmal Atrial Fibrillation after Catheter Ablation. Journal of Clinical Medicine, 12(5). DOI:10.3390/jcm12051933
[41] Wang, M., Chen, G., Luo, J., Fan, Z., Liu, Y., Xie, C., & Gong, Y. (2023). Case Report: Genetic profiling of small intestine metastasis from poorly differentiated non-small cell lung cancer: report of 2 cases and literature review of the past 5 years. Frontiers in Oncology, 13. DOI:10.3389/fonc.2023.1265749
[42] Wang, L., Mei, X., Zhou, Y., Zeng, J., Yang, T., Liao, L., Xiong, M., Zhao, X., & He, R. (2023). Abnormalities of the Serum Proteomic in Thrombosis After Central Venous Catheter (CVC) Insertion in Patients with End-stage Kidney Disease. Iranian Journal of Kidney Diseases, 17(6), 335–347. DOI:10.52547/ijkd.7671
[43] Ferreira, R. M. (2023). New evidence-based practice: Artificial intelligence as a barrier breaker. World Journal of Methodology, 13(5), 384–389. DOI:10.5662/wjm.v13.i5.384
[44] Satapathy, P., Padhi, B. K., & Al Khathlan, N. (2023). Artificial intelligence in respiratory therapy: Opportunities and ethical challenges. Respiratory Medicine, 217(000), 2. DOI:10.1016/j.rmed.2023.107342
[45] Cederberg, K. L. J., Sikes, E. M., & Mignot, E. (2024). Perceptions of exercise and restless legs syndrome: Results from a nationwide survey. Journal of Sleep Research, 33(1), e13980. DOI:10.1111/jsr.13980
[46] Shibata, Y., Kawayama, T., Muro, S., & Sugiura, H. (2023). Management goals and stable phase management of patients with chronic obstructive pulmonary disease in the Japanese respiratory society guideline for the management of chronic obstructive pulmonary disease 2022 (6th edition). Respiratory Investigation, 61(6), 773–780. DOI:10.1016/j.resinv.2023.08.007
[47] Wadgaonkar, M., Aggarwal, A., & Solanki, G. S. (2023). Asthma management with the help of digital therapeutics: Review of clinical trials. Indian Journal of Allergy, Asthma and Immunology, 37(1), 3–7. DOI:10.4103/ijaai.ijaai_21_23
[48] Galderisi, S. (2023). Personalisation of the management of schizophrenia and other primary psychoses. European Psychiatry, 66(S1), S45–S45. DOI:10.1192/j.eurpsy.2023.164
[49] Sue, G. R. (2023). Artificial Intelligence for Plastic Surgeons. Plastic and Reconstructive Surgery - Global Open, 11(6), E5057. DOI:10.1097/GOX.0000000000005057
[50] Połomska, J., Dydak, P., Sozańska, B., & Sikorska-Szaflik, H. (2023). Peanut Allergy and Component-Resolved Diagnostics Possibilities—What Are the Benefits? Nutrients, 15(24). DOI:10.3390/nu15245132
[51] Zhao, X., Meng, Q., Zhou, M., Luo, J., & Hu, L. (2023). Optimal treatment strategy and prognostic analysis for patients with non-metastatic pT4 colon adenocarcinoma. Frontiers in Oncology, 13(000), 14. DOI:10.3389/fonc.2023.1342289
[52] Assi, T., Khoury, R., Ibrahim, R., Baz, M., Ibrahim, T., & LE Cesne, A. (2023). Overview of the role of liquid biopsy in cancer management. Translational Oncology, 34. DOI:10.1016/j.tranon.2023.101702
[53] Rezaee, R., Parsa, Z., Ahmadzadeh, L., Ahmadian, L., Avazzadeh, S., & Marzaleh, M. A. (2023). Self-help application for obsessive-compulsive disorder based on exposure and response prevention technique with prototype design and usability evaluation: A cross-sectional study. Health Science Reports, 6(9). DOI:10.1002/hsr2.1577
[54] Riedell, P. A., Moses-Okwei, E., Sullivan, S., Carter, J., & Heggen, C. (2023). Accelerating the Adoption of New and Emerging Therapies for DLBCL: A Community-Based Quality Improvement Initiative. Blood, 142(Supplement 1), 7264–7264. DOI:10.1182/blood-2023-173686
[55] Krisanapan, P., Tangpanithandee, S., Thongprayoon, C., Pattharanitima, P., & Cheungpasitporn, W. (2023). Revolutionizing Chronic Kidney Disease Management with Machine Learning and Artificial Intelligence. Journal of Clinical Medicine, 12(8), 4. DOI:10.3390/jcm12083018

[56] Geng, J., Guo, Y., Xie, M., Li, Z., Wang, P., Zhu, D., Li, J., & Cui, X. (2023). Characteristics of endoplasmic reticulum stress in colorectal cancer for predicting prognosis and developing treatment options. Cancer Medicine, 12(10), 12000–12017. DOI:10.1002/cam4.5874

[57] Goktas, P., Karakaya, G., Kalyoncu, A. F., & Damadoglu, E. (2023). Artificial Intelligence Chatbots in Allergy and Immunology Practice: Where Have We Been and Where Are We Going? Journal of Allergy and Clinical Immunology: In Practice, 11(9), 2697–2700. DOI:10.1016/j.jaip.2023.05.042

[58] Huang, Y., Huang, X., Wang, A., Chen, Q., Chen, G., Ye, J., Wang, Y., Qin, Z., & Xu, K. (2023). Individualized treatment decision model for inoperable elderly esophageal squamous cell carcinoma based on multi-modal data fusion. BMC Medical Informatics and Decision Making, 23(1). DOI:10.1186/s12911-023-02339-5

[59] Shao, J., Feng, J., Li, J., Liang, S., Li, W., & Wang, C. (2023). Novel tools for early diagnosis and precision treatment based on artificial intelligence. Chinese Medical Journal Pulmonary and Critical Care Medicine, 1(3), 148–160. DOI:10.1016/j.pccm.2023.05.001

[60] Abramova, M. Y., Medveditskova, A. I., & Lukina, G. I. (2022). Personalized approach to the treatment of patients with dental anomalies with the involvement of dentists of related specialties. Russian Stomatology, 15(1), 20–24. DOI:10.17116/rosstomat20221501120

[61] Patra, I., Sugumar, D., Yuvaraj, S. A., Bari, M. A., Nuhmani, S., & Unnisa, A. (2023). Personalized behavioural psychotherapy and treatment for children with neurodegenerative disease using machine learning approaches. AIP Conference Proceedings, 2603. DOI:10.1063/5.0126160

[62] Sadigh, G., & Abi-Jaoudeh, N. (2022). Surgery Versus Ablation for Small Hepatocellular Carcinoma Treatment and Role of Personalized Medicine. Journal of the American College of Radiology, 19(11), 1224–1225. DOI:10.1016/j.jacr.2022.09.023

[63] Chalkias, A. (2022). Current Challenges in the Management of Sepsis and Septic Shock: Personalized, Physiology-Guided Treatment. Journal of Clinical Medicine, 11(15). DOI:10.3390/jcm11154565

[64] Navest, R., van Pelt, V., van de Lindt, T., Janssen, T., Nowee, M., Sonke, J., van der Heide, U., & van Houdt, P. (2022). PD-0896 Personalized MRI contrast for treatment guidance on an MRI-linac in patients with liver metastases. Radiotherapy and Oncology, 170, S788–S790. DOI:10.1016/s0167-8140(22)02975-9

[65] Zaikov, S. V., Bogomolov, A. Y., Katilov, O. V., Gumeniuk, G. L., Yakovenko, O. K., & Dmitrieva, E. M. (2023). Global Trends in Therapy of Patients with Bronchial Asthma — a Personalized Approach to Treatment. Asthma and Allergy, 2023(2), 35–46. DOI:10.31655/2307-3373-2023-2-35-46

[66] Mohamed, S. M., Naguib, E. A., Hamama, H., & Fawzi, E. M. M. (2022). Effect of personalized caries treatment protocol versus standard of care on salivary flow rate of high caries risk patients. International Journal of Health Sciences, 2567–2574. DOI:10.53730/ijhs.v6ns4.7412

[67] Oh, S. H., Park, J., Lee, S. J., Kang, S., & Mo, J. (2022). Reinforcement learning-based expanded personalized diabetes treatment recommendation using South Korean electronic health records. Expert Systems with Applications, 206. DOI:10.1016/j.eswa.2022.117932

[68] Lu, Y., & Du, X. (2023). Personalized prevention and treatment strategies for hypotension patients during maintenance hemodialysis. MEDS Clinical Medicine, 4(1). DOI:10.23977/medsc.2023.040110

[69] Opalinska, M., Kamiński, G., Dedecjus, M., Kowalska, A., Kolodziej, M., Saracyn, M., Garnuszek, P., Lenda-Tracz, W., Borkowska, A., Gąsior-Perczak, D., Budzyńska, A., Kubik, A., Kacperski, K., Szubstarska, P., Chalewska, W., Długosińska, J., Januszkiewicz-Caulier, J., Hubalewska-Dydejczyk, A., & Mikolajczak, R. (2023). DUONEN multicenter study - personalized PRRT treatment with 177Lu- or 177Lu/90Y-DOTA-TATE in patients with neuroendocrine tumors based on individual dosimetry. Endocrine Abstracts. DOI:10.1530/endoabs.90.p440

[70] Hagan, K. E. (2023). Unlocking the Potential of Effective Connectivity as a Neuroimaging Biomarker: A Path Toward Personalized Depression Treatment in Youth. Biological Psychiatry, 94(12), 910–912. DOI:10.1016/j.biopsych.2023.09.010

[71] Kaya, İ., Yakar, H., & Kesen, E. (2023). Low-Cost 3-D-Printer–Assisted Personalized Cranioplasty Treatment: A Case Series of 14 Consecutive Patients. World Neurosurgery, 175, e1197–e1209. DOI:10.1016/j.wneu.2023.04.098

[72] Crea, F. (2022). Personalized antithrombotic treatment of cardiovascular disease: a glimpse into the future based on current knowledge. European Heart Journal, 43(10), 925–929. DOI:10.1093/eurheartj/ehac086

[73] Wei, M., & Chu, C. Q. (2022). Prediction of treatment response: Personalized medicine in the management of rheumatoid arthritis. Best Practice and Research: Clinical Rheumatology, 36(1), 101741. DOI:10.1016/j.berh.2021.101741

[74] Zhu, J., Ye, J., Dong, L., Ma, X., Tang, N., Xu, P., Jin, W., Li, R., Yang, G., & Lai, X. (2023). Non-invasive prediction of overall survival time for glioblastoma multiforme patients based on multimodal MRI radiomics. International Journal of Imaging Systems and Technology, 33(4), 1261–1274. DOI:10.1002/ima.22869

[75] Molnar, M. J., & Molnar, V. (2023). AI-based tools for the diagnosis and treatment of rare neurological disorders. Nature Reviews Neurology, 19(8), 455–456. DOI:10.1038/s41582-023-00841-y

[76] Alshehri, F. S. (2023). The complex regional pain syndrome: Diagnosis and management strategies. Neurosciences (Riyadh, Saudi Arabia), 28(4), 211–219. DOI:10.17712/nsj.2023.4.20230034

[77] Shi, H. Y., Li, C. H., Chen, Y. C., Chiu, C. C., Lee, H. H., & Hou, M. F. (2023). Quality of life and cost-effectiveness of different breast cancer surgery procedures: a Markov decision tree-based approach in the framework of Predictive, Preventive, and Personalized Medicine. EPMA Journal, 14(3), 457–475. DOI:10.1007/s13167-023-00326-4

[78] Preskorn, S. H., & Rode, R. (2023). Personalized Medicine in the Treatment of a Patient with Obsessive-Compulsive Disorder with Clomipramine. Journal of Psychiatric Practice, 29(6), 469–475. DOI:10.1097/PRA.0000000000000750

[79] Weiskittel, T. M., Ung, C. Y., Correia, C., Zhang, C., & Li, H. (2022). De novo individualized disease modules reveal the synthetic penetrance of genes and inform personalized treatment regimens. Genome Research, 32(1), 124–134. DOI:10.1101/GR.275889.121

[80] Vromans, R. D., Pauws, S. C., van de Poll-Franse, L. V., & Krahmer, E. J. (2023). Effects of comparative information when communicating personalized risks of treatment outcomes: an experimental study. Journal of Risk Research, 26(3), 324–343. DOI:10.1080/13669877.2022.2128392

[81] VV, L. (2019). Personalized Approach in the Diagnostic and Treatment of Tinnitus. American Journal of Biomedical Science & Research, 3(5), 439–441. DOI:10.34297/ajbsr.2019.03.000712