A Literature Review on the Application of Reinforcement Learning in Healthcare Management

Valentina Popolo*, Teresa Angela Trunfio, Arianna Scala*****

** Department of Management Engineering, University Unipegaso, Naples, Italy (Corresponding author: valentina.popolo@unipegaso.it)*

*** Department of Advanced Biomedical Sciences, University Federico II, Naples, Italy (teresaangela.trunfio@unina.it)*

**** Department of Public Health, University Federico II, Naples, Italy (arianna.scala@unina.it)*

Abstract: Reinforcement Learning (RL) is a branch of Machine Learning and more generally of Artificial Intelligence used for complex systems. In particular, through a learning-by-trying and error method and a rewarding system, it is able to understand its environment and solve complicated decision-making tasks. Due to its ability, the RL has gained attention in various fields, including healthcare. Personalized Treatment Plans, Clinical Decision Support Systems and Resource Allocation in Hospitals are just a few examples that show the potential of RL not only in clinical practice but also in management. In this paper, we aim to perform a synthesis of the applications of RL to Healthcare Management (HM) by conducting a literature review of indexed scientific articles published in the last three years (2021-2023) on the topic. The purpose is to answer the research question on the success and critical factors in the implementation of RL to HM and assess its ability to manage a highly complex supply chain and resource system such as that of healthcare. Subsequently, a list of the most widely used algorithms and their performance was obtained, also discussing the ethical and legal implications that might limit their implementation. This review will add knowledge to the existing literature, especially providing neophytes with a preliminary document to consult before approaching this topic in its reality.

Keywords: Review, Machine Learning, Reinforcement Learning, Healthcare, Healthcare Management.

1. Introduction

Reinforcement learning (RL) is a branch of Machine Learning (ML) that closely mirrors the learning processes of humans and animals (Sutton and Barto, 2018). Many RL algorithms are directly inspired by biological behavior. In particular, an agent interacts with its environment by performing specific actions and perceiving environmental states from which it learns the correct behavior through the input of a positive feedback signal (Esteso et al., 2023). The typical steps in an RL algorithm include: (i) observing the environment; (ii) selecting a strategy; (iii) acting on the choice made; (iv) receiving a reward or penalty; (v) learning from the feedback received and changing strategy; (vi) repeating these steps until the optimal strategy is reached (Afsaneh et al., 2022). Each algorithm thus operates in two main phases: learning and inference. There are three distinct learning approaches in RL: Policy-based. Value-based, and Model-based, as illustrated in Figure 1.

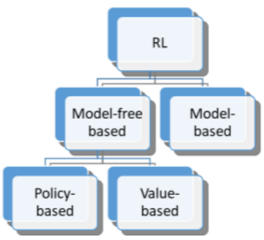

Fig. 1: Overview of RL.

Model-based algorithms learn a model of the environment that predicts the state of the environment after performing a particular action. An example of this is Monte Carlo Trees. In contrast, model-free algorithms focus on learns a policy (policy-model) that defines the relationship between states and actions. The policy specifies the probability of taking each action in a specific state, with policy gradients being a common example. Finally, value-based algorithms learn value function that maps states to expected rewards, which is used to estimate the reward for taking a certain action in a given state. Q-learning is one of the most popular value-based algorithms. Depending on the specific problem to be modeled, one of these types described above will be used. If the goal is to maximize the obtained reward, value-based algorithms will be employed. If the objective is to learn a policy that achieves a certain result, policy-based algorithms will be used. Finally, if the environment is stochastic or not easily observable model-based ones will be the choice. Due to its high potential, interest in RL has grown significantly in recent years. Figure 2 shows the trend of Google searches on the topic from 2004 until March 2024 ("Google Trends," n.d.).

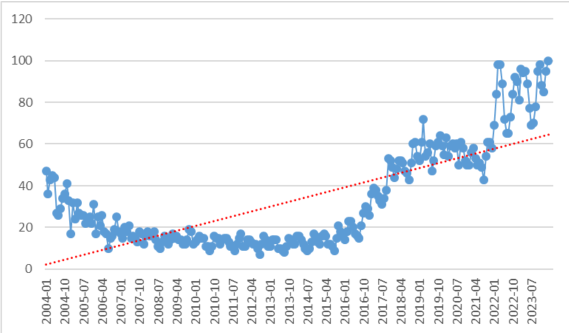

Fig. 2: Graph produced from data extracted from Google Trends on Google searches using the keywords "Reinforcement Learning" over the period 2004-2024.

The graph shows the time on the x-axis and the search interest on the y-axis. This is determined with respect to the highest point recorded. A value of 100 indicates the highest search frequency of the term, 50 indicates half of the searches. A score of 0, on the other hand, indicates that not enough data were found. Compared to the red trend line in Figure 2, in recent years there has been a wider variability and stronger growth than expected from the observation of previous years. RL, in fact, is not new but was theorised over 70 years ago in its simplest trial-and-error form. Only the technological and scientific advances of the last few years, however, have allowed for continued progress to more high-performance approaches such as deep learning.

1.1 Aim of work

Research is being published at an ever-increasing pace, with both the number of scientific publications and new online journals rapidly growing. Conducting literature reviews has thus becomes strategic, as they allow for the summarization of existing knowledge on a specific topic, clarification of controversies and identification of gaps and new insights (Michels and Schmoch, 2012). Although traditional literature reviews have limitations, e.g. selection bias and challenges in papers identification, they remain popular and valuable tool (Haddaway et al., 2015). In recent decades, the application of various ML techniques has expanded in the field of medical informatics for both diagnostic purposes (Loperto et al., 2022; Scala et al., 2022) and process optimization (Ponsiglione et al., 2023; Scala et al., 2023). ML techniques offer diverse options for data processing, making them an important focus for academics and researchers. Therefore, given the growing interest in this topic, particularly in this specific application context, summarizing this new body of knowledge is essential, especially to support junior researchers (Hasan and Bao, 2021). The purpose of this work is to investigate the implementation of RL algorithms in healthcare, with specific focus on Healthcare Management (HM).

2. Research methodology

As mentioned in the previous section, we conducted a literature search for scientific publications that apply to support HM. From the initial evaluation, it became evident that there are still relatively few contributions on this topic. In fact, most of the available research applies RL algorithms to healthcare or management as separate industrial component. Other studies combine the two keywords, but focus on disease management, such as diabetes, rather than HM. Figure 3 illustrates the methodology used for this review.

Literature Review Methodology	
Definition of research question	
RQ1 – What are the main application of RL in HM?	
RQ2 – What are the main algorithms for RL in HM?	
Selection of database	
Scopus, PubMed	
Search Terms	**Results**
Reinforcement Learning, Healthcare Management	11515 Publications
Inclusion criteria	**Results**
1 – Year of publications: [2022; 2024]	2866 Publications
2 – Languages: English	2861 Publications
3 – Article Type: Research articles	1575 Publications

Fig. 3: Literature review methodology.

As shown in Figure 3, two research questions and a single reference database were defined. We decided to focus on the last five years, i.e. how much scientific and technological progress, together with the availability of data, has enabled growth in the use of these approaches. Of the 1575 publications obtained after screening, title and abstract were analyzed in order to identify actual relevance to the subject matter. The articles found also include articles that deal separately with Machine Learning, Management and Healthcare. For this reason, only those articles that implement RL on the subject of Healthcare were selected by reading the titles and keywords in the abstract (60 articles). From these, articles were excluded that are inherent to clinical practice, working on healthcare data such as images, for the detection of diseases or alterations. At the end of this further screening, 40 articles were selected. The results were finally grouped into topics and it was decided to treat only one article per type as an example, not including all the contexts in which the same solution was used. In the end, 29 articles were collected and will be briefly presented in the next section divided into three main topics: Supply Chain, Resource Management, Patient and Disease Management and Wearable Sensor Management.

3. Results

This section presents the results of the literature search, divided into sub-sections based on specific topics.

3.1 Supply Chain

Following the analysis process, five articles were identified that examined the use of RL algorithms to optimize the supply chain. A summary of these articles is presented in Table 1.

In their article, Long et al. propose a sustainable supply chain development using artificial intelligence. The authors define various techniques and identify the best choice, verifying the model through simulation. The model based on the DDPG algorithm closely aligns with the target path, showing minimal error, indicating its effectiveness. The authors conclude by demonstrating how these tools can assist the healthcare supply chain in making intelligent decisions and selecting the best operational modes. However, the study's limitations include limited data collection and a narrow area of interest.

In contrast, Kitchat et al. use the DDPG algorithm to manage the supply chain during the complex period of the COVID-19 pandemic. They develop a robust system for allocating epidemic prevention materials in Taiwan pharmacies. Specifically, they discretize the number of masks and define a vector representing the supply for each pharmacy as an output. Using real data over four months and simulations with increasing mask demand, they show that as the number of masks increases, the model's performance improves in terms of Root Mean Square Error (RMSE) and reward.

Also focusing on COVID-19, Rey et al. address the distribution of vaccines using a model that considers different geographical areas, disease spread and a shared budget mechanism. The authors demonstrate that a RL strategy promoting global cooperation is the best solution, although they note limitations such as the unavailability of reliable and accurate data from some nations and the model used to simulate disease spread.

Ahmadi et al. use a similar algorithm to model inventory policies in a network of several regional hospitals and a central warehouse. They demonstrate that their approach, compared to traditional policy, offers higher speed, a lower probability of product shortages and a better patients service.

Finally, Saha et al. propose a stochastic semi-Markov decision process model solved by a multi-agent reinforcement learning method for managing drug inventory within a hospital. The model is validated with data from a multi-specialty hospital in India, showing that it not only improves inventory management but also enhances service delivery and reduces costs.

Table 1: Publications on Supply Chain and RL.

Publication	RL algorithm
(Long et al., 2023)	Depth Deterministic Policy Gradient (DDPG)
(Kitchat et al., 2024)	Deep Deterministic Policy Gradient
(Rey et al., 2023)	Thompson Sampling
(Ahmadi et al., 2022)	Q-learning, Deep Q-learning
(Saha and Rathore, 2024)	Multi-Agent

3.2 Resource Management

Another topic addressed in the literature is resource management. The first group of articles, presented in Table 2, concerns the allocation of resources.

Table 2: Publications on Resource Allocation and RL.

Publication	RL algorithm
(Talaat, 2022a)	Deep RL
(Talaat, 2022b)	
(Lazebnik, 2023)	Policy-based model
(Zong and Luo, 2022)	Multi-agent recurrent attention actor-critic
(Jangra and Mangla, 2023)	Q-Learning - SARSA
(Palani and Rameshbabu, 2024)	Deep RL
(Su et al., 2024)	Online Offloading Deep RL
(Zhang et al., 2023)	Q-Learning - BWACO

In the first two studies by Talaat, a new Effective Resource Allocation Strategy (ERAS) for the Fog environment is presented, which utilizes an RL algorithm in the allocation module. The proposed system operates on three levels. In first levels involves data acquisition from the patient through a series of sensors. When critical data is detected (e.g. an abnormality in the cardiac tracing), the staff is alerted. The second level focuses on resource allocation with low latency and appropriateness for handling the critical event. The third level features an effective prediction module using Probabilistic Neural Networks to predict a target field (e.g. the probability of a heart attack), based on one or more predictors. In his second study, Talaat optimizes the hyper-parameters of the algorithm and tests this optimized version by monitoring quality indicators. The results show a correlation between the selected hyper-parameters and the model's efficiency.

Lazebnik also employs an RL algorithm for resource allocation, using real-world data for training. This approach results in a solution that improves patient outcomes and cost-effectiveness compared to previous algorithms. Despite the model's validity and generalizability - according to the author's tests - imitations include not accounting for potential resource shortages due to illnesses or voluntary departures, variability in payment, and the role of support staff who, although not directly involved with patients, are essential for the hospital's proper functioning.

As in the previous section, specific studies are available for resource allocation during COVID-19 era, as demonstrated by Zong and Luo. Their model differs by focusing on optimal lockdown resource allocation rather than hospital resource allocation. It starts with a pandemic spread model between different states, predicting interactions between populations based on age and economic conditions.

Jangra and Mangla implement a variant of Q-learning, called SARSA, aiming to optimize resource scheduling in a cloud-based healthcare environment. The cloud layer processes requests with very low latency, enabling prompt intervention by healthcare personnel when necessary. The authors also highlight numerous potential causes of algorithm errors. Palani et al. propose a cryptography-based architecture to enhance security and reduce execution times by utilizing under-utilized resources. The simulated model results show optimal working time, resource allocation, and security in e-health systems. Su et al. address real-time task offloading in an intelligent healthcare network with dynamically varying environments, achieving reduced system task processing costs.

Finally, Zhang et al. apply these tools in a rapidly developing context, i.e. territorial care and in particular home care. They construct a network where the nodes are the homes of the elderly and the arcs represent the horizontal and vertical distances nurses travel. The goal is to minimize waiting times and ensure all requests are processed. However, the limitation lies in the proposed mathematical model, which is based on consultations with field experts rather than objective data.

Within the broader framework of Resource Management, the works shown in Table 3 are also included.

Table 3: Publications on Resource Management and RL.

Publication	RL algorithm
(Chen and Li, 2024)	Deep Q-network
(Guerrero et al., 2022)	Model-based policy optimization
(Zhong et al., 2024)	Deep RL, Kalman filtering
(Shuvo et al., 2023)	Multi-Objective RL
(Lakhan et al., 2023)	Multi-Agent Dueling Double Deep Q-Network
(Al-Marridi et al., 2024)	Q-Learning - DRLBTS
(Mishra et al., 2023)	Distributed RL

Chen et al. propose a dynamic system for teleconsultation scheduling to optimize start time. Specialists are the key resource for this strategy, with working hours and observed intervals between teleconsultations serving as model constraints. However, collaborations between clinical departments and the preferences of specialists and users are not considered, which are also constraints of this work. Guerrero et al. develop a decision-support system to validate building facilities according to current regulations in Spain. The model verifies parameters such as room ventilation, heating, lighting, hydraulics, gas distribution, layout of doors, windows, etc. The system is also intended for validating health building projects created by students during their training. Zhong et al. propose an RL approach for cloud computing, which has garnered significant interest in healthcare services due to its ability to provide various medical services via the Internet. Their method aims to optimize these services by improving availability, reliability, energy consumption, and response time. Similarly focused on service delivery, Shuvo et al. propose a multi-objective RL model to optimize the planning and expansion of hospital services, even in critical situations such as a pandemic. The model was tested on data from Florida and adapted to that specific context, demonstrating a reduction in service denial and an increase in costs.

In the last three articles, the Blockchain-based healthcare system is addressed. In the first article, a pioneering healthcare system is proposed for data sharing among all healthcare facilities and the ministry, balancing latency, security and cost. An RL algorithm is suggested to manage the system's heterogeneity. The second paper introduces a Deep PL-Aware Blockchain-based Task Scheduling algorithm that provides security and efficient scheduling for healthcare applications. The latest work by Mishra et al. presents a system called CogniSec, which employs a decentralized cognitive blockchain and RL architecture to address security issues in medical cyber-physical systems.

3.3 Patient and Disease Management

This section collects papers focusing on the clinical aspect of patient management including treatment and disease management, and concludes with articles modelling the spread of infections.

Table 4: Publications on Patient Management and RL.

Publication	RL algorithm
(Oh et al., 2022)	Deep Q-learning
(Li et al., 2022)	Deep Q-learning
(Raheb et al., 2022)	Deep Q-learning
(Emerson et al., 2023)	Q-learning
(Khalilpourazari and Hashemi Doulabi, 2022)	Q-learning
(Zeng et al., 2023)	Multi-reward Deep RL

Oh et al. propose a model based on RL algorithms to identify the best treatment for patients with type 2 diabetes. The approach involves creating clusters of patients with similar characteristics using k-means algorithm, thereby advancing the concept of personalized medicine. The model incorporates the patient's biomedical data (BMI, blood pressure, glycated hemoglobin levels, etc.) and assigns a unique metformin therapy according to the protocol. If the patient's condition worsens, the therapy dosage may be adjusted. The model was validated by comparing the simulated data with the expert medical prescriptions, demonstrating a reduction in complications and lower glycated hemoglobin levels. Using a similar RL algorithm, Li et al. develop a system to optimize sequential treatment for diseases such as sepsis and diabetes, focusing specifically on patients with diabetic ketoacidosis. The system recommends drug dosages for various physiological states using data from electronic medical records. Unlike other studies, their model employs a different agent structure at each time-step and an improved reward system. Raheb et al. and Emerson et al. also address diabetes from a clinical perspective. In the first paper, offline RL techniques are used to create glucose control policies for patients with type 1 diabetes. The algorithm adapts to the patient's initial data, updating insulin requirements over time. However, the authors note limitations, including the model's failure to account for stress and activity levels, and the lack of validation with real patient data. Further studies and developments are needed before clinical application. In the second paper, on the other hand, the authors propose an automated model that calculates insulin doses in clinical units for subcutaneous injection in patient with type 2 diabetes. Their model accurately emulates the basal-bolus process by accounting for both insulin infusion and absorption

The last two articles in Table 4 use RL algorithms to predict the course of the COVID-19 pandemic using the most up-to-date mathematical models. These studies also assess the potential impact of by governments actions.

3.4 Wearable Sensor Management

This final section focuses on the implementation of RL algorithms for managing wearable sensors (WS).

In the first study identified, Tripathy et al. explore the use of RL to adapt and record user information collected with wearable sensors (WS), tailoring the data according to specific environmental situations. In contrast, Arikumar et al. employ RL to tag data from wearable devices without storing all the data in the cloud, instead leveraging peripheral components. Their classification model achieved an accuracy of 99.67% but may be vulnerable to privacy issues. Finally, Li et al. propose an RL system for Online Mobile Charging Scheduling with optimal Quality of Sensing Coverage, resulting in a model that, after extensive simulations, proves to be stable and superior to existing algorithms.

Table 5: Publications on WS Management and RL.

Publication	RL algorithm
(Tripathy et al., 2024)	Q-learning
(Arikumar et al., 2022)	Deep RL
(Li et al., 2024)	Deep Q-Network

4. Discussion and Conclusion

In this paper, we present a synthetic literature review to identify the implementations of RL algorithms in the context of HM.

From an initial selection of over 1500 scientific articles in English published between 2022 and 2024, we selected 29 articles, categorizing them into four different topics: Supply Chain, Resource Management, Patient and Disease Management and Wearable Sensor Management. For each topic, we focused on presenting distinct types of work, avoiding redundant discussions across different application contexts.

Supply Chain Management and Resource Allocation are the most frequently implemented RL techniques. Additionally, the topic of block chain, is also highly valued by scientists. Numerous articles also explore the topic of Patient and Disease Management, aiming to identify the most effective treatment or supporting continuous monitoring of vital parameters. However, this topic, this topic tends to be closer to clinical practice rather than top-level management. The same applies to Wearable Sensor Management, where the primary objective is to manage devices and collected data useful for medical treatment.

In summary, while there is considerable literature on the application of RL to support patient diagnosis, treatment, and monitoring, the HM topic remains underexplored, especially concerning processes and care pathways. As shown by (Assadullah, 2019), various fields including clinical decision-making, risk assessment, care processes, continuity of care, coordination of care, safety of care, and managerial processes in healthcare, have been analyzed by Artificial Intelligence (AI). For instance, (Grant and McParland, 2019) demonstrate the effectiveness of AI in an emergency room using an e-triage model for categorizing patient severity. A search for "Artificial Intelligence, Healthcare Management" in the same databases yields more than 8,000 articles, including those by (Senapati et al., 2024)) on AI algorithms supporting the supply chain and (Mizan and Taghipour, 2022) on machine learning predicting patient arrival and optimizing resource allocation. Despite these developments, RL remains poorly implemented in HM, with areas such as clinical risk and process optimization still unaddressed, suggesting potential avenues for further research.

This work has limitations, including the small number of identified and discussed papers, the limited analysis period, and the restricted number of reference databases. Additionally, several keywords were not tested, which might have affected the comprehensiveness of the review.

References

Afsaneh, E., Sharifdini, A., Ghazzaghi, H., Ghobadi, M.Z., 2022. Recent applications of machine learning and deep learning models in the prediction, diagnosis, and management of diabetes: a comprehensive review. Diabetol. Metab. Syndr. 14, 196. https://doi.org/10.1186/s13098-022-00969-9

Ahmadi, E., Mosadegh, H., Maihami, R., Ghalehkhondabi, I., Sun, M., Süer, G.A., 2022. Intelligent inventory management approaches for perishable pharmaceutical products in a healthcare supply chain. Comput. Oper. Res. 147, 105968. https://doi.org/10.1016/j.cor.2022.105968

Al-Marridi, A.Z., Mohamed, A., Erbad, A., 2024. Optimized blockchain-based healthcare framework empowered by mixed multi-agent reinforcement learning. J. Netw. Comput. Appl.

224, 103834. https://doi.org/10.1016/j.jnca.2024.103834

Arikumar, K.S., Prathiba, S.B., Alazab, M., Gadekallu, T.R., Pandya, S., Khan, J.M., Moorthy, R.S., 2022. FL-PMI: Federated Learning-Based Person Movement Identification through Wearable Devices in Smart Healthcare Systems. Sensors 22, 1377. https://doi.org/10.3390/s22041377

Assadullah, M.M., 2019. Barriers to Artificial Intelligence Adoption in Healthcare Management: A Systematic Review. https://doi.org/10.2139/ssrn.3530598

Chen, W., Li, J., 2024. Teleconsultation dynamic scheduling with a deep reinforcement learning approach. Artif. Intell. Med. 149, 102806. https://doi.org/10.1016/j.artmed.2024.102806

Emerson, H., Guy, M., McConville, R., 2023. Offline reinforcement learning for safer blood glucose control in people with type 1 diabetes. J. Biomed. Inform. 142, 104376. https://doi.org/10.1016/j.jbi.2023.104376

Esteso, A., Peidro, D., Mula, J., Díaz-Madroñero, M., 2023. Reinforcement learning applied to production planning and control. Int. J. Prod. Res. 61, 5772–5789. https://doi.org/10.1080/00207543.2022.2104180

Google Trends [WWW Document], n.d. . Google Trends. URL https://trends.google.it/trends/explore?date=all&q=Reinforcement%20Learning&hl=it (accessed 3.29.24).

Grant, K., McParland, A., 2019. Applications of artificial intelligence in emergency medicine 96.

Guerrero, J.I., Miró-Amarante, G., Martín, A., 2022. Decision support system in health care building design based on case-based reasoning and reinforcement learning. Expert Syst. Appl. 187, 116037. https://doi.org/10.1016/j.eswa.2021.116037

Haddaway, N. r., Woodcock, P., Macura, B., Collins, A., 2015. Making literature reviews more reliable through application of lessons from systematic reviews. Conserv. Biol. 29, 1596–1605. https://doi.org/10.1111/cobi.12541

Hasan, N., Bao, Y., 2021. Understanding current states of machine learning approaches in medical informatics: a systematic literature review. Health Technol. 11, 471–482. https://doi.org/10.1007/s12553-021-00538-6

Jangra, A., Mangla, N., 2023. An efficient load balancing framework for deploying resource schedulingin cloud based communication in healthcare. Meas. Sens. 25, 100584. https://doi.org/10.1016/j.measen.2022.100584

Khalilpourazari, S., Hashemi Doulabi, H., 2022. Designing a hybrid reinforcement learning based algorithm with application in prediction of the COVID-19 pandemic in Quebec. Ann. Oper. Res. 312, 1261–1305. https://doi.org/10.1007/s10479-020-03871-7

Kitchat, K., Lin, M.-H., Chen, H.-S., Sun, M.-T., Sakai, K., Ku, W.-S., Surasak, T., 2024. A deep reinforcement learning system for the allocation of epidemic prevention materials based on DDPG. Expert Syst. Appl. 242, 122763. https://doi.org/10.1016/j.eswa.2023.122763

Lakhan, A., Mohammed, M.A., Nedoma, J., Martinek, R., Tiwari, P., Kumar, N., 2023. DRLBTS: deep reinforcement learning-aware blockchain-based healthcare system. Sci. Rep. 13, 4124. https://doi.org/10.1038/s41598-023-29170-2

Lazebnik, T., 2023. Data-driven hospitals staff and resources allocation using agent-based simulation and deep reinforcement learning. Eng. Appl. Artif. Intell. 126, 106783. https://doi.org/10.1016/j.engappai.2023.106783

Li, J., Wang, H., Jiang, C., Xiao, W., 2024. A deep reinforcement learning approach for online mobile charging scheduling with optimal quality of sensing coverage in wireless rechargeable sensor networks. Ad Hoc Netw. 156, 103431. https://doi.org/10.1016/j.adhoc.2024.103431

Li, T., Wang, Z., Lu, W., Zhang, Q., Li, D., 2022. Electronic health records based reinforcement learning for treatment optimizing. Inf. Syst. 104, 101878. https://doi.org/10.1016/j.is.2021.101878

Long, P., Lu, L., Chen, Q., Chen, Y., Li, C., Luo, X., 2023. Intelligent selection of healthcare supply chain mode – an applied research based on artificial intelligence. Front. Public Health 11. https://doi.org/10.3389/fpubh.2023.1310016

Loperto, I., Scala, A., Rossano, L., Carrano, R., Federico, S., Triassi, M., Improta, G., 2022. Use of regression models to predict glomerular filtration rate in kidney transplanted patients, in: 2021 International Symposium on Biomedical Engineering and Computational Biology, BECB 2021. Association for Computing Machinery, New York, NY, USA, pp. 1–4. https://doi.org/10.1145/3502060.3503627

Michels, C., Schmoch, U., 2012. The growth of science and database coverage. Scientometrics 93, 831–846. https://doi.org/10.1007/s11192-012-0732-7

Mishra, S., Chakraborty, S., Sahoo, K.S., Bilal, M., 2023. Cogni-Sec: A secure cognitive enabled distributed reinforcement learning model for medical cyber–physical system. Internet Things 24, 100978. https://doi.org/10.1016/j.iot.2023.100978

Mizan, T., Taghipour, S., 2022. Medical resource allocation planning by integrating machine learning and optimization models. Artif. Intell. Med. 134, 102430. https://doi.org/10.1016/j.artmed.2022.102430

Oh, S.H., Lee, S.J., Park, J., 2022. Effective data-driven precision medicine by cluster-applied deep reinforcement learning. Knowl.-Based Syst. 256, 109877. https://doi.org/10.1016/j.knosys.2022.109877

Palani, S., Rameshbabu, K., 2024. A secured energy aware resource allocation and task scheduling based on improved cuckoo search algorithm and deep reinforcement learning for e-healthcare applications. Meas. Sens. 31, 100988. https://doi.org/10.1016/j.measen.2023.100988

Ponsiglione, A.M., Trunfio, T.A., Amato, F., Improta, G., 2023. Predictive Analysis of Hospital Stay after Caesarean Section: A Single-Center Study. Bioengineering 10, 440. https://doi.org/10.3390/bioengineering10040440

Raheb, M.A., Niazmand, V.R., Eqra, N., Vatankhah, R., 2022. Subcutaneous insulin administration by deep reinforcement learning for blood glucose level control of type-2 diabetic patients. Comput. Biol. Med. 148, 105860. https://doi.org/10.1016/j.compbiomed.2022.105860

Rey, D., Hammad, A.W., Saberi, M., 2023. Vaccine allocation policy optimization and budget sharing mechanism using reinforcement learning. Omega 115, 102783. https://doi.org/10.1016/j.omega.2022.102783

Saha, E., Rathore, P., 2024. A smart inventory management system with medication demand dependencies in a hospital supply chain: A multi-agent reinforcement learning approach. Comput. Ind. Eng. 191, 110165. https://doi.org/10.1016/j.cie.2024.110165

Scala, A., Loperto, I., Triassi, M., Improta, G., 2022. Risk Factors Analysis of Surgical Infection Using Artificial Intelligence: A Single Center Study. Int. J. Environ. Res. Public. Health 19, 10021. https://doi.org/10.3390/ijerph191610021

Scala, A., Trunfio, T.A., Improta, G., 2023. Classification and regression model to manage the hospitalization for laparoscopic cholecystectomy. Sci. Rep. 13, 14700. https://doi.org/10.1038/s41598-023-41597-1

Senapati, T., Sarkar, A., Chen, G., 2024. Enhancing healthcare supply chain management through artificial intelligence-driven group decision-making with Sugeno–Weber triangular norms in a dual hesitant q-rung orthopair fuzzy context. Eng. Appl. Artif. Intell. 135, 108794. https://doi.org/10.1016/j.engappai.2024.108794

Shuvo, S.S., Symum, H., Ahmed, M.R., Yilmaz, Y., Zayas-Castro, J.L., 2023. Multi-Objective Reinforcement Learning Based Healthcare Expansion Planning Considering Pandemic Events. IEEE J. Biomed. Health Inform. 27, 2760–2770. https://doi.org/10.1109/JBHI.2022.3187950

Su, X., Fang, X., Cheng, Z., Gong, Z., Choi, C., 2024. Deep reinforcement learning based latency-energy minimization in smart healthcare network. Digit. Commun. Netw. https://doi.org/10.1016/j.dcan.2024.06.008

Sutton, R.S., Barto, A.G., 2018. Reinforcement Learning, second edition: An Introduction. MIT Press.

Talaat, F.M., 2022a. Effective prediction and resource allocation method (EPRAM) in fog computing environment for smart healthcare system. Multimed. Tools Appl. 81, 8235–8258. https://doi.org/10.1007/s11042-022-12223-5

Talaat, F.M., 2022b. Effective deep Q-networks (EDQN) strategy for resource allocation based on optimized reinforcement learning algorithm. Multimed. Tools Appl. 81, 39945–39961. https://doi.org/10.1007/s11042-022-13000-0

Tripathy, J., Balasubramani, M., Rajan, V.A., S, V., Aeron, A., Arora, M., 2024. Reinforcement learning for optimizing real-time interventions and personalized feedback using wearable sensors. Meas. Sens. 33, 101151. https://doi.org/10.1016/j.measen.2024.101151

Zeng, J.-Y., Lu, P., Wei, Y., Chen, X., Lin, K.-B., 2023. Deep reinforcement learning based medical supplies dispatching model for major infectious diseases: Case study of COVID-19. Oper. Res. Perspect. 11, 100293. https://doi.org/10.1016/j.orp.2023.100293

Zhang, T., Liu, Y., Yang, X., Chen, J., Huang, J., 2023. Home health care routing and scheduling in densely populated communities considering complex human behaviours. Comput. Ind. Eng. 182, 109332. https://doi.org/10.1016/j.cie.2023.109332

Zhong, C., Darbandi, M., Nassr, M., Latifian, A., Hosseinzadeh, M., Jafari Navimipour, N., 2024. A new cloud-based method for composition of healthcare services using deep reinforcement learning and Kalman filtering. Comput. Biol. Med. 172, 108152. https://doi.org/10.1016/j.compbiomed.2024.108152

Zong, K., Luo, C., 2022. Reinforcement learning based framework for COVID-19 resource allocation. Comput. Ind. Eng. 167, 107960. https://doi.org/10.1016/j.cie.2022.107960