

REVIEW ARTICLE



Prediction of chemotherapy-related complications in pediatric oncology patients: artificial intelligence and machine learning implementations

Seda Ardahan Sevgili¹✉ and Selmin Şenol²

© The Author(s), under exclusive licence to the International Pediatric Research Foundation, Inc 2022

Although the overall incidence of pediatric oncological diseases tends to increase over the years, it is among the rare diseases of the pediatric population. The diagnosis, treatment, and healthcare management of this group of diseases are important. Prevention of treatment-related complications is vital for patients, particularly in the pediatric population. Nowadays, the use of artificial intelligence and machine learning technologies in the management of oncological diseases is becoming increasingly important. With the advancement of software technologies, improvements have been made in the early diagnosis of risk groups in oncological diseases, in radiology, pathology, and imaging technologies, in cancer staging and management. In addition, these technologies can be used to predict the outcome in chemotherapy treatment of oncological diseases. In this context, this study identifies artificial intelligence and machine learning methods used in the prediction of complications due to chemotherapeutic agents used in childhood cancer treatment. For this purpose, the concepts of artificial intelligence and machine learning are explained in this review. A general framework for the use of machine learning in healthcare and pediatric oncology has been drawn and examples of studies conducted on this topic in pediatric oncology have been given.

Pediatric Research (2023) 93:390–395; <https://doi.org/10.1038/s41390-022-02356-6>

IMPACT:

- Artificial intelligence and machine learning are advanced tools that can be used to predict chemotherapy-related complications.
- Algorithms can assist clinicians' decision-making processes in the management of complications.
- Although studies are using these methods, there is a need to increase the number of studies on artificial intelligence applications in pediatric clinics.

INTRODUCTION

The National Cancer Institute defines cancer as a disease that occurs when the cell grows uncontrollably and invades other parts of the body.¹ Cancer incidence rates in the pediatric population have increased by an average of 0.8% per year since 1975, depending on the type of cancer among both children and adolescents. Despite the increase in incidence, thanks to advances in cancer treatment and technology, mortality rates are decreasing in both children and adolescents. Thus, according to the data of 2019, mortality in children was calculated as 1.8 per 100,000 cases and in adolescents as 2.8 per 100,000 cases.²

The most commonly used method in cancer is treatment with chemotherapeutic agents. It can be used alone or along with surgical treatment and radiotherapy, depending on the clinical condition of the patient.³ The main purpose of chemotherapy is to destroy cancer cells, stop their spread, and thus cure the disease. It should be applied to increase the quality of life by eliminating the symptoms related to the disease (palliative), to reduce the risk of relapse after surgery or radiotherapy (adjuvant), or to

facilitate treatment by applying it before surgery or radiotherapy (neoadjuvant).^{4–6}

Cancer treatment is based on the type and stage of cancer. Some factors should be considered when creating a treatment regimen to be administered to people diagnosed with cancer. Individuals are divided into different risk groups, considering the age at diagnosis, tumor stage, pathological findings related to cell cycle and differentiation, and genetic factors such as oncogenic mutations or translocations.⁷ For example, in hematological malignancies, the higher the risk group of individuals, the higher the dose and frequency of the treatment they receive. As the dose of treatment increases, whether it is radiation therapy or chemotherapy, it is inevitable that treatment-related complications will occur.⁸ Chemotherapy not only affects cancer cells but also has effects on healthy cells. In particular, it has serious effects on the gastrointestinal, hematopoietic, reproductive, and nervous systems.⁹ The most common gastrointestinal complications are nausea, vomiting, oral-anal mucositis, bowel changes, dysphagia, taste changes, and weight loss.^{10–12} Bleeding due to

¹Pediatric Nursing Department, Ege University, Faculty of Nursing, 35100 Bornova-Izmir, Turkey. ²Pediatric Nursing Department, Kütahya Health Sciences University, Faculty of Health Sciences, 43100 Kütahya, Turkey. ✉email: sedardahan@gmail.com

Received: 14 March 2022 Revised: 8 October 2022 Accepted: 12 October 2022

Published online: 27 October 2022

thrombocytopenia, infection, and fever due to neutropenia, dyspnea, and fatigue due to anemia are hematopoietic complications.^{13–16} Also, complications such as organ toxicity, sensory disorders, cognitive disabilities, pain, anxiety, and depression are also observed.^{17,18} In a population with so many treatment-related complications, the importance of complication management is indisputable. With today's technology, there have been great developments in the health services for oncological diseases. One of these technological developments is the integration of artificial intelligence (AI) and machine learning (ML) into the field, which is the main subject of this review. The purpose of this review is to provide examples of studies using AI and ML technologies, to explain the methods and algorithms that can be used on the subject, and to shed light on the authors for future studies.

Artificial intelligence

AI is a subfield of computer science that studies algorithms to perform human-like cognitive functions such as learning and problem-solving.¹⁹ Just like human cognitive abilities, AI uses prior knowledge to analyze patterns within data. The biggest difference between AI and human cognition is that AI can perform these analyses in a very short time and without any physical or mental effort.²⁰ The program used for AI is the intelligent agent that interacts with the environment. The agent determines the state of the environment through sensors or inputs and can influence the state through actuators. Translation of inputs into actuators occurs through functions within the agent, which is called the agent's control policy. In other words, learning algorithms are used to imitate human intelligence in AI, whose main purpose is to capture human-like intelligence in machines.²¹

Today, AI applications have been developed in almost all fields that interact with information technologies (especially health, military, transportation, manufacturing industry, agriculture, trade, etc.) and this development continues. For instance, speech assistants such as Alexa, Siri, Cortana, or Google Assistant, designed to be almost indistinguishable from humans, a self-driving and maneuverable car, aircraft, or ship, a consultant on online commerce sites or mobile services, recommending various products and obtaining bank loans. Algorithms developed for validation are some intelligent systems developed for AI with which humans interact directly. This growth and commercialization of AI applications are especially the success of ML algorithms.²²

Machine learning

ML was defined by Arthur Samuel as a subfield of computer science that gives "computers the ability to learn without being explicitly programmed".²³ Developed from pattern recognition and computational learning theories in AI, ML includes the creation and examination of algorithms that can learn from data and make predictions. ML algorithms are dynamic and tend to evolve as more data is provided, in other words, learning.²⁴ ML algorithms build a model from sample inputs (training dataset) and make data-based predictions and decisions.²⁵ In this respect, ML is a critical field of AI as it can model based on experiences and help predict future events accurately. ML has very successful applications in a wide variety of fields such as health, education, sensor technology, transportation, manufacturing, and finance.²⁶

ML techniques are classified according to input/label variables (supervised, unsupervised, and reinforcement learning) and output/feature variables (handcraft-based and non-handcraft-based techniques).²⁴ Supervised algorithms predict a known outcome. This learning is used to perform one of the two tasks: classification and regression. The main difference between these two tasks is that the label or output predicted in classification is discrete or nominal (categorical), whereas in regression the predicted is continuous (numerical) or ordered value.²⁷ Some of the most commonly used methods in supervised learning are Naive Bayes

(NB), linear and logistic regression (LR), support vector machines (SVM), neural networks (NN), decision tree (DT), k-nearest neighborhood (KNN) and Random Forests (RF).²⁶ As aforementioned, supervised learning is often used to predict future events from past data. For instance, if the objective is to predict the mortality of the patients after acute lymphoblastic leukemia (ALL) treatment, the model will be trained on a dataset containing information on patient, disease, and treatment-related complications as well as the associated outcome (survival) for each individual. It is usually determined by measuring the predictive power on datasets that are excluded in the model development process (test and validation data) to measure the performance of models developed through supervised learning.²⁷

Unsupervised learning is an ML method used to detect relationships and dependencies in unlabeled data, hence this method is not intended to predict an outcome. Some of the most used methods in this type of learning are clustering, association rule mining, anomaly (outlier) detection, density estimation, and representation learning. The main purpose of unsupervised learning is to induce the internal data structure to produce a useful representation without the aid of explicit class labels.^{24,26}

Reinforcement learning can be characterized as a group of algorithms that work sequentially. A reinforcement algorithm, which is characterized as an agent at each step, predicts a feature in a future step based on past features. The algorithm learns which actions will reach the maximum score over a given amount of time by penalizing the correct output for the incorrect output. Thus, the aim of learning the best policy is achieved.^{24,25,28}

The categories of features/outputs are divided into handcrafted and non-handcrafted. The handcraft-based method is based on extracting an unlimited number of open explicit features from the dataset. The features are often shaped by information that decision makers or experts think interacts with their goals and are subject to qualitative judgment. Traditional ML algorithms are generally used in the handcraft-based method. However, in non-handcrafted feature-based techniques, raw data are processed as part of the learning. The algorithm learns and then identifies suitable features to reduce prediction error or improve classification performance. Large datasets are needed for high performance in this method, so the human interpretation may not be necessary. The handcraft-based method works with deep learning techniques.^{29,30}

Use of machine learning in healthcare

The integration of ML applications into medicine, nursing, and other health sciences started with the adoption of electronic medical records (EMR), abandoning the tradition of writing patient data on printed documents in clinics.³¹ The ability of applications to analyze large amounts of data facilitates the integration of medical and healthcare services. Due to the increasing amount of big digital data, studies involving ML applications have also gained momentum. Thanks to today's technology, the power of software programs and computer informatics also contribute to this increase.^{32,33} There are examples of ML, particularly in genomics,³⁴ pharmacology,^{35,36} radiology, and other imaging technologies.

The widespread use of AI applications in oncology is due to some characteristics of this disease group. First, many people are diagnosed with cancer every year in the world, especially in the adult population, depending on the type of cancer, mortality rates are generally high. According to cancer statistics prepared by the American Cancer Society, it is estimated that 1.9 million people will be diagnosed with cancer in the United States alone in 2022, and 609,360 people will experience cancer-related mortality.² Second is the cost of cancer treatment. According to the Annual Report to the Nation on the Status of Cancer published by the National Cancer Institute in 2021; the cancer-related economic burden was predicted to be \$21.1 billion for 2019.³⁷ The third

feature is that both the planning and management of cancer treatment require a very comprehensive information synthesis. Cancer requires a long-term treatment process, and a substantial amount of data must be collected and processed even for only one individual during this period.²⁰ Evaluating this whole process with traditional methods without ML is not only costly but also imposes a burden in terms of time and effort.

The use of ML in the field of oncology has been frequently encountered recently.^{38–40} Whether it is pathology or radiology, the use of ML for oncological imaging methods is common. In pathology, ML is used to diagnose and predict treatment responses based on pathological image patterns, as well as tumor detection and segmentation.^{41–44} In radiology, ML can improve different aspects of medical imaging.^{45,46} It is used in tumor identification, evaluation of new anomalies, and response to treatment by comparing the tumor with previous results. It also uses ML applications in obtaining radiological images, creating and interpreting reports, comparing radiological findings with different clinical data, and making clinical decisions.⁴⁷ To sum up, ML applications in oncological diseases are used for diagnosis and early detection of the disease, cancer classification, staging, predicting, and evaluating treatment response.

Literature review

As described earlier in this review, ML can be used to improve diagnostic accuracy and predict treatment outcomes by learning from available data. The purpose of this review is to shed light on the literature on the prediction of chemotherapy-related complications in the pediatric oncology population with ML applications^{14,15,17,48,49} In this part, in addition to the chemotherapy-related complications; cancer susceptibility⁵⁰ and cancer survival⁵¹ key sections and different study examples are also included. Although the number of studies on the mentioned subject has increased, as far as we know, it has not reached the desired level. A limited number of articles related to the subject could be reached in the literature review and the details of these articles were explained in the literature review. Furthermore, Table 1 contains the general scheme of these works, divided into themes and key sections.

To mathematically model the myelosuppression caused by chemotherapy, biological data should be presented to understand the effects of chemotherapeutic drugs on the bone marrow. In this context, in the study by Cuplov and Andre,¹⁴ conducted with 24 children diagnosed with rhabdomyosarcoma, estimation of chemotherapy-related hematological toxicity was performed. In the study, the effects of a specific chemotherapy protocol on blood results (platelets, neutrophils) of pediatric patients were estimated. All the children were selected from the children who underwent the IVA (ifosfamide, actinomycin D, and vincristine) protocol for treating rhabdomyosarcoma. The gradient boosting regression parameters were used as an ML algorithm in this study. In the study, the results of predicted and observed blood values were also compared to test the validation of the model.

Zhan et al. conducted a study with 249 newly diagnosed ALL patients in the 1–18 age group. The authors stated that the study aimed to predict the risk of neutropenia and fever developing after high-dose methotrexate (HD-MTX) infusion in B cell-ALL patients using an AI approach. Patients in the standard and intermediate-risk groups according to two different chemotherapy protocols were included in the dataset according to the MTX dose they received (2 or 5 g/m²). In the study, neutropenia and fever were evaluated by examining the EMRs of the patients at the end of the HD-MTX infusion until the next HD-MTX course or 14 days after the HD-MTX infusion. An absolute neutrophil count value less than $0.5 \times 10^9/L$ is considered severe neutropenia; body temperature $>38.3^\circ C$ was considered a fever. A total of 521 patients have received multiple infusions of MTX; data in the study were divided into two as training and test sets. The authors used five different

modeling algorithms: RF, SVM, NB, classification and regression tree, and DT. In the study, the highest accuracy value of HD-MTX-related neutropenia and fever was reached with the RF algorithm (AUC: 0.927 and 0.870, respectively). At the end of the study, the authors stated that the models were validated to predict the risk of neutropenia and fever with ML applications. They concluded that the models could be helpful in clinical decision making for oncologists.¹⁵

In the study by Naushad et al.,⁴⁸ the hematological toxicity risk of 6-mercaptopurine (6-MP), a purine analog used in the maintenance treatment of pediatric patients with ALL, was evaluated using ML algorithms. In the study, the data of 96 pediatric patients who were in the maintenance phase were examined. The dependent variable of the study is the total leukocyte count (TLC) in the hemogram results of the children at regular intervals. The decrease in TLC examined on the 43rd, 71st, and 99th days of maintenance was considered to be hematological toxicity. The authors used the Construction of Classification and Regression Tree algorithm in the study. The authors stated an accuracy value of 93.6% and the AUC value of 0.9649 for this model. It was concluded that the model could help tailor the dose of 6-MP therapy because of the high accuracy estimation of toxicity.⁴⁸

Some chemotherapeutic agents have serious toxicity complications on the body organs. Especially, the cardiotoxic effect of anthracyclines is known. Anthracycline cardiotoxicity may result in left ventricular (LV) dysfunction and heart failure and may lead to the development of cardiomyopathy even years after the last dose of chemotherapy. A study by Chaix et al.⁴⁹ studied 289 childhood cancer survivors 3 years after the last anthracycline exposure. A nested case-control design was used in the study; despite a low dose ($\leq 250 \text{ mg/m}^2$) of doxorubicin exposure, 183 cases were selected as extreme phenotype due to reduced LV ejection fraction. 106 control subjects were selected as the extreme phenotype due to preserved LV ejection fraction despite a high dose ($>250 \text{ mg/m}^2$) of doxorubicin. Study variables are classified as clinical (age of anthracycline exposure, sex, ethnicity), treatment (anthracycline dose, chest radiation), use of cardioprotective dexrazoxane, and biological (genes) predictors. In the study, the authors developed a risk prediction model that includes genetic and clinical predictors using the RF algorithm. A total of three models have been developed: the clinical model, the genetic model, and the model containing both genetic and clinical information. The combined RF model was noted to outperform clinical and genetic RF models with a higher AUC (0.71) higher specificity, higher positive predictive value, and lower misclassification rate. The model with the lowest performance was found to be clinical RF (AUC = 0.59).⁴⁹

Lu et al.¹⁷ conducted a cross-sectional study with children, adolescents, and caregivers in the 8–17 age group who had survived cancer. The authors stated the purpose of the study as testing the validity of natural language processing (NLP) and ML algorithms in determining the pain and fatigue symptoms experienced by patients due to cancer treatment. After the patient-reported outcome data obtained from the in-depth interviews were written down, two different NLP (Bidirectional Encoder Representations from Transformers (BERT) and Word2Vec) and ML (SVM or Extreme Gradient Boosting) methods were used to extract the semantic features of the sentences suitable for analysis and to ensure their validation. In the study's results, the authors stated that BERT showed higher accuracy in both symptoms with 0.931 (95% CI 0.905–0.957) when compared to Word2Vec/SVM and Word2Vec/Extreme Gradient Boosting. They also noted that the study used standard metrics to test the validity of their NLP/ML models, including precision, sensitivity, specificity, accuracy, F1 score, areas under the ROC curve, and AUC.¹⁷

In the future, ML applications could be used primarily for cancer prevention, especially in the era of big data oncology. Pan et al.⁵⁰

Table 1. Examples of studies involving machine learning methods in pediatric oncology.

Theme	Authors	Cancer type	Problem	No. of patients	Input (variables)	Output (labels)	Key section	Used model/s
Active therapy process	Naushad et al. (2019) ⁴⁸	Leukemia	6-MP-related leucopenia	96	Dose of chemotherapeutic agent	Laboratory values (total leukocyte count)	Cancer therapy-related acute toxicity	CART
	Cuplov and Andre (2020) ¹⁴	RMS	Ifosfamide-related hematological toxicities	24	Chemotherapy protocol (ifosfamide-induced hematological toxicity)	Laboratory values (platelet and neutrophil counts)		GBR
	Zhan et al. (2021) ¹⁵	Leukemia	MTX-related neutropenia and fever	521	Dose of chemotherapeutic agent	Laboratory values (absolute neutrophil count and body temperature)		RF, SVM, NB, DT CART
Therapy result	Lu et al. (2021) ¹⁷	Leukemia, HL, CNS/non-CNS tumors	Therapy-related pain and fatigue	52	Patient-reported pain and fatigue symptoms	Pain and fatigue symptoms experienced by patients		BERT, SVM, XGBoost
	Chaix et al. (2020) ⁴⁹	Leukemia, sarcomas, blastomas, NHL, HL, Wilms tumor	Anthracycline-related cardiotoxicity	289	Clinical predictors Treatment predictors Biological predictors	Contribution of rare and low-frequency single-nucleotide variant	Cancer therapy-related long-term toxicity	RF
	Pan et al. (2017) ⁵⁰	Leukemia	Leukemia recurrence	570	Patient and disease characteristics	Relapse in newly diagnosed pediatric ALL patients	Cancer susceptibility	RF
	Kashef et al. (2020) ⁵¹	Leukemia	Survival	241	Demographic characteristics Medical information Treatment-related complications	Survival	Cancer survival	DT, SVM, LDA, MLR, XGBoost, RF

RF random forest, SVM support vector machine, DT decision tree, NB Naive Bayes, XGBoost Extreme Gradient Boosting, GBR gradient boosting regression, LDA linear discriminant analysis, MLR multinomial linear regression, CART classification and regression tree, BERT Bidirectional Encoder Representations from Transformers, 6-MP 6-mercaptopurine, MTX methotrexate, NHL non-Hodgkin lymphoma, HL Hodgkin lymphoma, RMS rhabdomyosarcoma, CNS central nervous system.

conducted a study to predict the risk of recurrence in children with ALL using ML applications. The sample of the study consists of 570 newly diagnosed children with ALL. In the study, WBC count of the child, age at diagnosis, sex, prednisone response, lymphoblastic cell percentage in bone marrow aspiration performed on day 15 and day 33, minimal residual disease level, and risk group variables were used to estimate the risk of relapse. In the study, 336 data were used for training ML algorithms. Then, two different tests were carried out, with the remaining 224 data being 150 and 84 data. The algorithms used are specified as RF, SVM, LR, and DT algorithms. The highest accuracy value (ACC 0.831) was achieved with the RF algorithm. After the authors determined that RF was the algorithm that gave the highest accuracy in estimation, they used the same algorithm again to estimate the risk group of 84 children. The algorithm gave high accuracy in estimating relapse in different risk groups. Based on the results, the authors stated that the model has a good generalization ability to predict childhood ALL recurrence.⁵⁰

In Kashef et al.'s study,⁵¹ the authors stated that the main purpose of the study was to predict the survival of pediatric ALL patients based on the classification of clinical and medical data. In the study, the reports of 241 patients were examined and the demographic characteristics, medical information, and treatment-related complications of these patients were determined as independent variables. The dependent variable of the study is the treatment outcome, that is, the survival status of the children. In the aforementioned study, 144 data were grouped as training data and 96 data were grouped as test data, and two scenarios were designed for data analysis. The first scenario considered all pediatric ALL patients, whereas the second scenario excluded patients with an unknown cause of death. General classification algorithms (DT, SVM linear discriminant analysis, LR, Extreme Gradient Boosting, RF) were applied using RStudio, and comparisons were made to find the model with superior performance. Experimental results show that the Extreme Gradient Boosting algorithm outperforms the compared classifiers with an accuracy of 88.5% (95% CI: 82.3–94.0) in the originally designed scenario. The superior model in the second scenario is SVM with an accuracy of 94.90% (95% CI: 88.49–98.32). In addition, based on these results, the authors stated that the most determining factor for the treatment outcome is the frequency of fever, and the second factor is neutropenia.⁵¹

Although it is thought that ML applications will be used extensively in medical sciences in the future, there are some challenges and limitations that may be encountered in designing and applying ML algorithms. Data availability is a critical problem, especially since ML algorithms require large data for the performance and validity of the developed model. In addition, the quality of the data obtained is also very important. If the data are not recorded properly, the frequency of noisy data may be high in the dataset and the lost values may increase. This may lead to both a decrease in the number of data to be used in the model and an increase in the time required for the data pre-processing steps. Moreover, in terms of effectiveness, the developed model may not always need to be applied, it may be more appropriate to use traditional diagnostic methods in some cases.⁵² Another important challenge encountered in the application of ML in medical sciences is an ethical problem. Some of these problems can be listed as informed consent for data use, algorithmic fairness and biases, data privacy, surveillance, security, and transparency.⁵³

CONCLUSION

ML-based AI applications are used in the diagnosis and staging of cancer, estimation of treatment outcomes, and determination of prognosis. Medical procedures, which would be more troublesome in terms of time, cost, and effort with traditional diagnosis and

follow-up methods, can be transformed into a more practical process with ML applications. Despite all these facilitating aspects, there are also some limitations in the integration of ML into clinical settings: difficulties in creating quality and systematic big data for algorithms to achieve the most optimal results, problems in accessing the data they need, providing external validation of the model, reflecting the results of the application to the clinic and ethical dilemmas.

In this review, the use of ML-based AI applications, which have been gaining importance recently, in the chemotherapy treatment of oncological diseases has been tried to be conveyed. Although there are studies on the subject in the literature, it is seen that many studies focus on the diagnosis and staging of cancer. However, it is necessary to focus on post-treatment complications in oncological diseases where survival has reached high rates because of the development of cancer treatment in the last decades. At this point, there is a need to enrich the literature with studies focusing on the quality of life of individuals by improving their clinical conditions. In this context, we suggest that ML studies should focus on different complications that may develop after chemotherapy.

REFERENCES

1. National Cancer Institute (accessed 10 January 2022). <https://www.cancer.gov/about-cancer/understanding/what-is-cancer#definition> (2021).
2. American Cancer Society. Cancer facts & figures 2022 (accessed 8 August 2022). <https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2022.html#:~:text=The%20Facts%20%26%20Figures%20annual%20report,deaths%20in%20the%20United%20States> (2022).
3. Senapati, S., Mahanta, A. K., Kumar, S. & Maiti, P. Controlled drug delivery vehicles for cancer treatment and their performance. *Signal Transduct. Target Ther.* **3**, 1–19 (2018).
4. Basmadjian, R. B. et al. Developing a prediction model for pathologic complete response following neoadjuvant chemotherapy in breast cancer: a comparison of model building approaches. *JCO Clin. Cancer Inform.* **6**, 1–10 (2022).
5. Mao, N. et al. Intratumoral and peritumoral radiomics for preoperative prediction of neoadjuvant chemotherapy effect in breast cancer based on contrast-enhanced spectral mammography. *Eur. Radiol.* **32**, 3207–3219 (2022).
6. Teachey, D. T., Hunger, S. P. & Loh, M. L. Optimizing therapy in the modern age: differences in length of maintenance therapy in acute lymphoblastic leukemia. *Blood* **137**, 168–177 (2021).
7. Ramesh, S. et al. Applications of artificial intelligence in pediatric oncology: a systematic review. *JCO Clin. Cancer Inform.* **5**, 1208–1219 (2021).
8. Dupuis, L. L. et al. Optimizing symptom control in children and adolescents with cancer. *Pediatr. Res.* **86**, 573–578 (2019).
9. Chen, C. B. et al. Acute pancreatitis in children with acute lymphoblastic leukemia correlates with L-asparaginase dose intensity. *Pediatr. Res.* **92**, 459–465 (2022).
10. Song, Q., Zhang, J., Wu, Q., Li, G. & Leung, E. L. H. Kanglaite injection plus fluorouracil-based chemotherapy on the reduction of adverse effects and improvement of clinical effectiveness in patients with advanced malignant tumors of the digestive tract: a meta-analysis of 20 RCTs following the PRISMA guidelines. *Medicine* **99**, e19480 (2020). 17.
11. Lee, K. A. et al. The gut microbiome: what the oncologist ought to know. *Br. J. Cancer* **125**, 1197–1209 (2021).
12. Yao, Z. et al. Moxibustion for alleviating chemotherapy-induced gastrointestinal adverse effects: a systematic review of randomized controlled trials. *Complement Ther. Clin. Pract.* **46**, 1–11 (2022).
13. Nearing, J. T. et al. Infectious complications are associated with alterations in the gut microbiome in pediatric patients with acute lymphoblastic leukemia. *Front. Cell Infect. Microbiol.* **9**, 1–14 (2019).
14. Cuplov, V. & André, N. Machine learning approach to forecast chemotherapy-induced haematological toxicities in patients with rhabdomyosarcoma. *Cancers* **12**, 1944 (2020).
15. Zhan, M. et al. Machine learning to predict high-dose methotrexate-related neutropenia and fever in children with B-cell acute lymphoblastic leukemia. *Leuk. Lymphoma* **62**, 2502–2513 (2021).
16. Joseph, A., Joshua, J. M., & Mathews, S. M. Chemotherapy-induced neutropenia among breast cancer patients in a tertiary care hospital: risk and consequences. *J. Oncol. Pharm. Pract.* **0**, 1–5 (2022).
17. Lu, Z. et al. Natural language processing and machine learning methods to characterize unstructured patient-reported outcomes: validation study. *J. Med. Internet Res.* **23**, e26777 (2021).

18. Juan, Z. et al. Probiotic supplement attenuates chemotherapy-related cognitive impairment in patients with breast cancer: a randomised, double-blind, and placebo-controlled trial. *Eur. J. Cancer* **161**, 10–22 (2022).
19. Pethani, F. Promises and perils of artificial intelligence in dentistry. *Aust. Dent. J.* **66**, 124–135 (2021).
20. Chua, I. S. et al. Artificial intelligence in oncology: path to implementation. *Cancer Med.* **10**, 4138–4149 (2021).
21. Das, S., Dey, A., Pal, A. & Roy, N. Applications of artificial intelligence in machine learning: review and prospect. *Int. J. Comput. Appl.* **115**, 31–41 (2015).
22. Riedl, M. O. Human-centered artificial intelligence and machine learning. *Hum. Behav. Emerg. Tech.* **1**, 33–36 (2019).
23. Samuel, A. L. Some studies in machine learning using the game of checkers. *IBM J. Res. Dev.* **3**, 210–229 (1959).
24. Goldenberg, S. L., Nir, G. & Salcudean, S. E. A new era: artificial intelligence and machine learning in prostate cancer. *Nat. Rev. Urol.* **16**, 391–403 (2019).
25. Ongsulee, P. Artificial intelligence, machine learning and deep learning. 2017 15th International Conference on ICT and Knowledge Engineering 1–6 (2017).
26. Dogan, A. & Birant, D. Machine learning and data mining in manufacturing. *Expert Syst. Appl.* **166**, 114060 (2021).
27. Shouval, R., Fein, J. A., Savani, B., Mohty, M. & Nagler, A. Machine learning and artificial intelligence in haematology. *Br. J. Haematol.* **192**, 239–250 (2021).
28. Panch, T., Szolovits, P. & Atun, R. Artificial intelligence, machine learning and health systems. *J. Glob. Health* **8**, 1–8 (2018).
29. Hassan, N., Gomaa, W., Khoriba, G. & Haggag, M. Credibility detection in Twitter using word n-gram analysis and supervised machine learning techniques. *Int. J. Intell. Eng. Syst.* **13**, 291–300 (2020).
30. Saba, T. Computer vision for microscopic skin cancer diagnosis using handcrafted and non-handcrafted features. *Microsc. Res. Tech.* **84**, 1272–1283 (2021).
31. Bertsimas, D. & Wiberg, H. Machine learning in oncology: methods, applications, and challenges. *JCO Clin. Cancer Inform.* **4**, 885–894 (2020).
32. D'Amore, B., Smolinski-Zhao, S., Daye, D. & Uppot, R. N. Role of machine learning and artificial intelligence in interventional oncology. *Curr. Oncol. Rep.* **23**, 1–8 (2021).
33. Cuocolo, R., Caruso, M., Perillo, T., Ugga, L. & Petretta, M. Machine learning in oncology: a clinical appraisal. *Cancer Lett.* **481**, 55–62 (2020).
34. Xu, J. et al. Translating cancer genomics into precision medicine with artificial intelligence: applications, challenges and future perspectives. *Hum. Genet.* **138**, 109–124 (2019).
35. Liu, H. et al. Predicting effective drug combinations using gradient tree boosting based on features extracted from drug-protein heterogeneous network. *BMC Bioinforma.* **20**, 1–12 (2019).
36. Ekins, S. et al. Exploiting machine learning for end-to-end drug discovery and development. *Nat. Mater.* **18**, 435–441 (2019).
37. Yabroff, K. R. et al. Annual report to the nation on the status of cancer, part 2: patient economic burden associated with cancer care. *J. Natl. Cancer Inst.* **113**, 1670–1682 (2021).
38. Elemento, O., Leslie, C., Lundin, J. & Tourassi, G. Artificial intelligence in cancer research, diagnosis and therapy. *Nat. Rev. Cancer* **21**, 747–752 (2021).
39. Huynh, E. et al. Artificial intelligence in radiation oncology. *Nat. Rev. Clin. Oncol.* **17**, 771–781 (2020).
40. Daldrup-Link, H. Artificial intelligence applications for pediatric oncology imaging. *Pediatr. Radiol.* **49**, 1384–1390 (2019).
41. Niazi, M. K. K., Parwani, A. V. & Gurcan, M. N. Digital pathology and artificial intelligence. *Lancet Oncol.* **20**, e253–e261 (2019).
42. Lonche, V. Y. & Bhasin, B. Artificial intelligence and its potential in oncology. *Drug Disco. Today* **24**, 228–232 (2019).
43. Saito, A. et al. Prediction of early recurrence of hepatocellular carcinoma after resection using digital pathology images assessed by machine learning. *Mod. Pathol.* **34**, 417–425 (2021).
44. Ak, M., Toll, S. A., Hein, K. Z., Colen, R. R. & Khatua, S. Evolving role and translation of radiomics and radiogenomics in adult and pediatric neuro-oncology. *AJNR Am. J. Neuroradiol.* **43**, 792–801 (2022).
45. Hyun, S. H., Ahn, M. S., Koh, Y. W. & Lee, S. J. A machine-learning approach using PET-based radiomics to predict the histological subtypes of lung cancer. *Clin. Nucl. Med.* **44**, 956–960 (2019).
46. Dong, X. et al. Deep learning-based attenuation correction in the absence of structural information for whole-body positron emission tomography imaging. *Phys. Med. Biol.* **65**, 055011 (2020).
47. Sadaghiani, M. S., Rowe, S. P. & Sheikhabaehi, S. Applications of artificial intelligence in oncologic 18F-FDG PET/CT imaging: a systematic review. *Ann. Transl. Med.* **9**, 823–834 (2021).
48. Naushad, S. M. et al. Classification and regression tree-based prediction of 6-mercaptopurine-induced leucopenia grades in children with acute lymphoblastic leukemia. *Cancer Chemother. Pharmacol.* **83**, 875–880 (2019).
49. Chaix, M. A. et al. Machine learning identifies clinical and genetic factors associated with anthracycline cardiotoxicity in pediatric cancer survivors. *JACC CardioOncol.* **2**, 690–706 (2020).
50. Pan, L. et al. Machine learning applications for prediction of relapse in childhood acute lymphoblastic leukemia. *Sci. Rep.* **7**, 1–9 (2017).
51. Kashaf, A., Khatibi, T. & Mehrvar, A. Treatment outcome classification of pediatric acute lymphoblastic leukemia patients with clinical and medical data using machine learning: a case study at MAHAK hospital. *Inf. Med. Unlocked.* **20**, 1–15 (2020).
52. Rahmani, A. M. et al. Machine learning (ML) in medicine: review, applications, and challenges. *Mathematics* **9**, 2970 (2021).
53. Naik, N. et al. Legal and ethical consideration in artificial intelligence in health-care: who takes responsibility? *Front. Surg.* **9**, 1–6 (2022).

AUTHOR CONTRIBUTIONS

S.A.S. contributed to conception and design, drafting the article, and final approval of the version to be published. S.S. contributed to revising the article and final approval of the version to be published.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Correspondence and requests for materials should be addressed to Seda Ardahan Sevgili.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.