Osmangazi Tıp Dergisi
Osmangazi Journal of Medicine 2020; 42 (3): 339- 349

Review / Derleme

Machine Learning in Radiation Oncology

Radyasyon Onkolojisinde Makine Öğrenmesi

Melek Akcay, Durmus Etiz

Eskisehir Osmangazi University Faculty of Medicine, Department of Radiation Oncology
Eskisehir, Turkey

Abstract: Artificial intelligence (AI) is a computer science that tries to imitate human-like intelligence on machines using computer software and algorithms without direct human stimuli to perform certain tasks. Machine learning (ML) is the subunit of AI that uses data-driven algorithms that learn to emulate human behavior based on a previous example or experience. Deep learning (DL) is an ML technique that utilizes deep neural networks to construct a model. The growth and sharing of data, increased computing power, and developments in ML have initiated a transformation in healthcare. Advances in radiation oncology have generated substantial data that must be integrated with computed tomography (CT) imaging, dosimetry, and other imaging modalities before each fraction. There are many algorithms used in Radiation Oncology. Each of these methods has advantages and limitations and different computing requirements. In this paper, we aimed to review the radiotherapy (RT) process by identifying the specific areas in which the quality and efficiency of ML can be increased and a workflow chart can be created. The RT stage is divided into seven groups as patient assessment, simulation, contouring, planning, quality assessment (QA), treatment application, and patient follow-up. A systematic evaluation of the applicability, limitations and advantages of ML algorithms was performed at each stage.

Keywords: Radiotherapy, machine learning, deep learning, artificial intelligence

Özet: Yapay zeka (YZ), belirli görevleri yerine getirmek için doğrudan insan uyaranları olmadan bilgisayar yazılımı ve algoritmaları kullanan makinelerde insan benzeri zekayı taklit etmeye çalışan bir bilgisayar bilimidir. Makine öğrenimi (MÖ), önceki bir örneğe veya deneyime dayanarak insan davranışını taklit etmeyi öğrenen veri odaklı algoritmalar kullanan yapay zekanın alt birimidir. Derin öğrenme (DÖ), bir model oluşturmak için derin sinir ağlarını kullanan bir MÖ tekniğidir. Verilerin büyümesi ve paylaşımı, artan bilgi işlem gücü ve MÖ'deki gelişmeler sağlık hizmetlerinde bir dönüşüm başlatmıştır. Radyasyon onkolojisindeki ilerlemeler, her fraksiyon öncesi yapılan bilgisayarlı tomografi (BT) görüntülemesi, dozimetri ve görüntüleme ile entegre edilmesi gereken önemli miktarda veri üretmiştir. Radyasyon Onkolojisinde kullanılan birçok algoritma vardır. Bu yöntemlerin her birinin farklı hesaplama gücü gereksinimleriyle avantajları ve sınırlamaları vardır. Bu derlemede, radyoterapi (RT) sürecinin, MÖ ile kalitesinin ve verimliliğinin arttırılabileceği belirli alanları belirleyerek iş akışı sırası ile gözden geçirme amaçlanmıştır. RT aşaması, hasta değerlendirmesi, simülasyon, konturlama, planlama, kalite kontrol, tedavi uygulama ve hasta takibi olarak yedi gruba ayrılmıştır. Her aşamaya MÖ algoritmalarının uygulanabilirliği, sınırlamaları, avantajları ile ilgili sistematik bir değerlendirme yapılmıştır.

Anahtar Kelimeler: Radyoterapi, makine öğrenmesi, derin öğrenme, yapay zeka

ORCID ID of the authors: M.A 0000-0002-9042-9489, D.E 0000-0002-2225-0364

Received **1**9.02.2020 ***Accepted*** 20.02.2020 ***Online published*** 28.04.2020

Correspondence: Melek AKÇAY- Eskisehir Osmangazi University Faculty of Medicine, Department of Radiation Oncology, Eskisehir, Turkey.
e-mail: mcakcay@ogu.edu.tr

Cite this article as:
Akcay M, Etiz D, Machine Learning in Radiation Oncology, Osmangazi Journal of Medicine2020;42:(3):339-349
Doi: *10.20515/otd.691331*

1. Introduction

Artificial intelligence (AI) is a computer science that tries to imitate human-like intelligence on machines using computer software and algorithms without direct human stimuli to perform certain tasks (1,2). Machine learning (ML) is the subunit of AI that uses data-driven algorithms that learn to emulate human behavior based on a previous example or experience (3). Deep learning (DL) is an ML technique that utilizes deep neural networks to construct a model. Increased computing power and reduced financial obstacles have eventually led to the emergence of the DL era (4). Figure 1 shows a schematic representation that reflects the overlapping areas of and relationships between AI, ML, and DL.

AI: Artificial intelligence, ML: Machine learning, DL: Deep learning

Figure 1. Artificial intelligence

The growth and sharing of data, increased computing power, and developments in ML have initiated a transformation in healthcare. Advances in radiation oncology have generated substantial data that must be integrated with computed tomography (CT) imaging, dosimetry, and other imaging modalities before each fraction.

Evidence-based medicine relies on randomized controlled trials designed for a large patient population. However, increasing the number of clinical and biological parameters to be investigated makes it difficult to design research (5). New approaches are required for all patient populations. Clinicians should use all diagnostic tools, such as medical imaging, blood tests, and genetic tests to decide on the appropriate combination of treatments (radiotherapy, chemotherapy, targeted therapy, immunotherapy, etc.). There are a series of individual differences that are responsible for each patient's disease or associated with the treatment response and clinical outcome. The concept of personalized treatment is based on identifying and using these factors in each case (6). The integration of such large and heterogeneous data is a problem that needs to be overcome to produce accurate models. Lambin et al. described in detail the features that should be included and considered in a prediction model as follows (7):

- Clinical features; e.g., patient performance status, grade and stage of tumor, blood tests, and patient surveys,
- Treatment features; e.g., dose distribution and chemotherapy,
- Imaging features; e.g., tumor size and volume, metabolic uptake, and radiomics,
- Molecular features; e.g., intrinsic radiosensivity, hypoxia, proliferation, and normal tissue reactions.

Common algorithms used in radiation oncology:

- Decision trees in which a simple algorithm answers questions in a predetermined order to create classes that exclude each other (8),

- Naive Bayes classifiers (9), which create probabilistic dependencies between variables (9),
- K-nearest Neighbors, used for classification and regression, in which a feature is classified according to its closest neighbor in the dataset (10),
- Support Vector Machine (SVM), in which a trained model categorizes new data (11),
- Artificial Neural Networks (ANNs) inspired by biological neural networks (12),
- DL, a variant of ANNs using multiple layers of neurons (2).

Each of these methods has advantages and limitations and different computing requirements.

In this paper, we aimed to review the radiotherapy (RT) process by identifying the specific areas in which the quality and efficiency of ML can be increased and a workflow chart can be created. The RT stage is divided into seven groups as patient assessment, simulation, contouring, planning, quality assessment (QA), treatment application, and patient follow-up, as shown in the flow chart given in Figure 2. A systematic evaluation of the applicability, limitations and advantages of ML algorithms was performed at each stage.

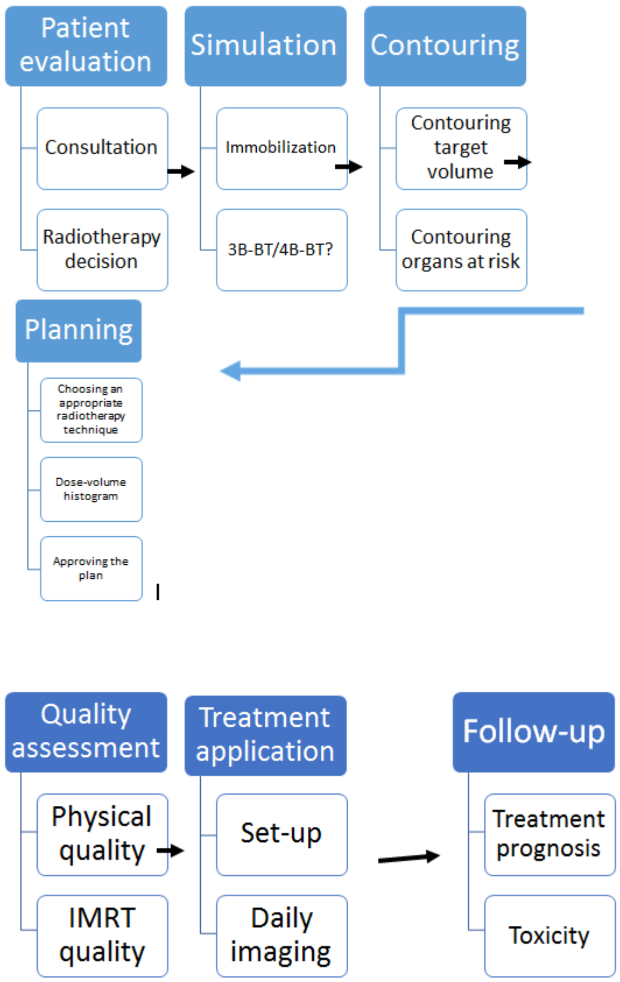

Figure 2. Radiotherapy workflow chart

Evaluation of Stages in Radiation Oncology by Machine Learning

Patient Assessment

The RT process begins with the patient's consultation, in which the radiation oncologist evaluates the risks and benefits of treatment and the clinical status of the patient to determine a treatment strategy. Tumor stage, mutation or genetic status (e.g., O6-methylguanine DNA methyltransferase - MGMT), viral status (e.g., HPV), previous and current treatments, surgical margin status, and general performance status are among the useful information to evaluate the potential benefit of treatment. At the same time, age, comorbidities, organ functions (e.g., kidney and liver), proximity between the tumor and normal critical tissues are parameters affecting tolerability to treatment. These are all features that can be used to construct treatment outcome and toxicity prediction models. These models can then be used to identify risks and benefits and guide doctors. Having sufficient information to define and classify the information available at this stage is important for the successful implementation of any prediction model (13).

The radiation oncologist should consider many factors during the evaluation of the patient, as well the interactions between these factors, and make a treatment decision accordingly. At this point, data-based prediction models can guide the doctor and make the decision-making process faster and more accurate. For example, when a patient with lung cancer is evaluated for stereotactic RT, the patient's respiratory functions, lung capacity, tumor size, proximity of the tumor to critical organs, comorbidities, and patient performance will affect both the treatment response and toxicity. If a model is constructed using these and other similar factors, the response and toxicity rates can be determined before starting treatment. As another example, in a patient that has undergone breast-sparing surgery after a left breast cancer diagnosis, a modeling including patient and treatment features can be established to predict whether this patient would benefit from the breath-hold technique.

Big data is needed to create these prediction models. The transition to the use of ML will increase the cooperation between health centers during the data collection phase and also help standardize treatments (14).

It has been suggested in various studies that logistic regression analysis and decision trees are equally effective in establishing a balance between the interpretability of the results and accurate predictions (15, 16) while for accuracy rates, random forest or gradient boosting and SVM algorithms are used (17,18).

Simulation

After the radiotherapy decision is made, a good simulation is required to select the correct treatment. The immobilization technique, scanning interval, and treatment area should be determined well. Preliminary preparations, such as whether fiducial use, full/empty bladder, and full/empty rectum are required for simulation, kidney function tests if intravenous contrast will be applied, and fasting status should be considered thoroughly. An accurate and good simulation is critical to obtaining a high-quality, robust treatment plan for the patient. In clinical practice, it is not uncommon to repeat CT due to deficiencies and inaccuracies during CT simulation, such as inadequate scanning interval, inadequate/incorrect immobilization technique, inappropriate level of bladder/rectum content, and hardware-related artifacts (14).

There are many questions that can be answered with ML algorithms to increase the overall workflow efficiency; e.g., Will this patient benefit from the use of an intravenous contrast agent? Which immobilization technique should be used? Can 4DCT be beneficial for this patient?

Contouring

In the standard workflow, the target volume and organs at risk (OAR) are manually contoured section by section by the radiation oncologist. As a result, this is a long process and causes a high degree of variability among contours, which constitutes one of the greatest sources of uncertainty in treatment planning (19).

Various attempts have been undertaken for automatic contouring, with the most common in clinical practice being atlas-based segmentation. First, the target image is mapped to one or more selected reference images. Then, the contours in the reference image are transferred to the target image (20). Atlas-based methods depend on the choice of atlas and the accuracy of reference images (21).

The ML approach in contouring directly learns the voxel structure of each image and combines it with previous information. Successful techniques include statistical and decision-learning classifiers, and DL has also recently been added to this list. Convolutional neural networks (CNNs) have been used for contouring organs at risk, and thoracic cancer, head and neck and prostate cancer (22-24). Table 1 presents the summary of the studies conducted. In a study by Lustberg et al., it was reported that ML and atlas-based contouring provided 61% and 22% time efficiency compared to manual contouring (22). This time saving is very important, especially when intensive clinics are considered.

Table 1. Machine learning in organ contouring

Publication	Tumor localization	Machine learning	Contouring	Patient number	Results
Lustberg et al., 2018 (22)	Lung	Deep learning, CNNs	Contoured by CT. OAR: *Lung *Esophagus *Heart *Mediasten * Spinal cord	20	*Manual contouring: 20 min *Atlas-based contouring: 7.8 min *Deep learning contouring: 10 min User adjustment of contours generated by machine learning reduced contouring time of OAR in lung radiotherapy.
Ibragimov and Xing, 2017 (23)	Head-neck	Deep learning, CNNs	Contoured by CT. OAR: * Spinal cord *Mandibula *Parotid gland *Submandibular gland *Larynx *Pharynx *Eyes *Optic nerves *Optic chiasm	50	Superior or comparable performance was achieved by CNN contouring of ms, mandibula, larynx, pharynx, eyes, and optic nerves compared to existing systems. CNN had low performance in the contouring of parotid and submandibular glands and optic chiasm; thus, additional imaging modalities are needed for the contouring of these organs.
Guo et al., 2016 (24)	Prostate	Deep learning	Contouring of the prostate by MRI.	66	Modern MRI achieved superior accuracy in prostate contouring compared to other methods.

CNNs, Convolutional neural networks; CT, Computed tomography; OAR, Organs at risk; MRI, Magnetic resonance imaging

Tumor volume contouring is often more difficult due to the different shape, size and localization of tumors, lack of clear boundaries, and dependence on the knowledge and experience of the oncologist. However, there are tumor contouring studies on brain, breast, oropharyngeal and rectal cancers (25-28), which are summarized and presented in Table 2.

Table 2. Machine leaning in target volume contouring

Publication	Tumor localization	Machine learning	Contouring	Patient number	Results
Men et al., 2018 (26)	Breast	Deep dilated residual network (DD-ResNet)	CTV	800	Two different deep learning methods were compared. The method proposed by the authors performed contouring in 15 sec per patient compared to 4 sec and 21 sec obtained from the other two DL methods. The DSC of the proposed method for contouring accuracy was 0.91.
Cardenas et al., 2018 (27)	Oropharynx	Deep learning (deep auto-encoders)	High-risk CTV	52	The median DSC for contouring accuracy was 0.81 (0.62-0.90). Automatic contouring can be undertaken to prevent variability in contouring data between clinicians.
Men et al., 2017, (28)	Rectum	Deep learning (deep dilated convolutional neural network)	CTV OAR: Bladder Right and left femur heads Small intestine Colon	278	The mean DSC values were 87.7 for CTV, 93.4 % for the bladder, 92.1% for the left femur head, 92.3% for the right femur head, 65.3% for the small intestine, and 61.8% for the colon.

CTV, Clinic target volume; OAR, Organs at risk; DSC, Dice similarity coefficient

Learning algorithms are trained to maximize the similarity measures between outputs and the samples provided. Therefore, although these algorithms are increasingly skilled at imitating human-drawn contours, they are limited by the quality of the training samples. It is considered that machines cannot be more ‘accurate' than the human input received as clinically essential facts, and their accuracy can only be meaningful in the context of individuals and institutional protocols until more concrete consensus definitions on threshold are specified (3).

Planning

The radiotherapy planning process is complicated. Failure in planning can cause life-threatening situations, such as missing the tumor or applying high doses of radiation to normal tissue. As the technology advances, the margin applied to the tumor decreases; thus, making it possible to miss a tumor even with a small margin of error.

After the target volumes and RAOs are defined, the planning process continues with the determination of targets and dosimetric targets for OAR, choosing an appropriate treatment technique (e.g., 3BKRT, IMRT, VMAT, and protons), achieving planning targets, and evaluating and approving the plan. Most ML practices focus on the stage of plan evaluation (29-31). Although related techniques are collectively called 'knowledge-based planning (KBP)', both current academic research and commercial products are limited to the prediction of dose-volume histograms (DVH) at accepted intervals (32-35). KBP methods develop fixed relationships from the geometrical and dosimetric parameters of previous plans. Rather than making a new start for each patient, this method draws on previous experience related to the initiation of optimization parameters in order to predict applicable DVH or voxel dose distributions and serve as personalized starting points for dosimetric changes (36-38). Plans created in this way are generally reported to successfully wrap around the target volume similar to manual plans and have better OAR doses (39).

The automatic generation of plans is also possible after the dosimetric targets are determined and the appropriate technique is selected. Studies have been conducted to solve various aspects of related problems; e.g., predicting the best beam directions (40, 41). It is considered that automatic treatment planning is suitable for the 'reinforcement learning' technique, in which the algorithm can make decisions in the light of certain rules and restrictions. Basically, the algorithm will make a decision (e.g., it will increase the weight of a particular constraint) and learn from the treatment planning system whether the decision is in the right direction. The difficulty of automatic planning using reinforcement learning is its requirement of integration with the treatment planning systems (14).

Quality Assessment

QA is very important for the evaluation of planned RT, detection of errors, and reporting. The features of the radiotherapy QA program allowing error detection and prevention, and QA of treatment devices are very suitable for the application of ML (42-45). Li and Chan developed an application to predict the performance of linear accelerators (Linac) over time (43). The daily QA of RT in cancer treatment closely monitors Linac performance and is critical for the continuous improvement of patient safety and quality of care. Cumulative quality assessment measurements are valuable to understand Linac behavior and allow medical physicists to identify trends in output and take preventive actions. In their study, Li and Chan applied ANNs and the autoregressive moving average time series prediction model to five years of Linac QA data. Then, they performed verification tests and other evaluations for all models and concluded that ANNs algorithm could be applied accurately and effectively to dosimetry and QA (43). Valdes et al. developed ML applications to estimate the QA transition rates of IMRT and detect problems in the Linac imaging system automatically (42). Carlson et al. developed an ML approach to predict multi-leaf collimator (MLC) position errors (46). Inconsistencies between the planned and transmitted movements of MLCs are a major source of error in dose distribution during RT. In the study of Valdes et al., for the prediction models of ML, various factors, such as leaf movement parameters, leaf position and speed, and the movement of the leaf toward or away from isocenter of MLC were determined from the plan files. The position differences between synchronized DICOM-RT planning files and DynaLog files reported during QA delivery were used for the training of the models. To assess the effect on the patient, the planned and predicted DVHs were compared with the DVH in the positions that were applied treatment. In all cases, the DVH parameters predicted for OAR, especially around the treatment area were found to be closer to the DVH in the position given treatment compared to the planned parameters (46).

Treatment Application

During radiation therapy, adjustments to treatment may be needed to ensure the proper implementation of the plan. The adjustments may be required as a result of both online factors, such as the patient’s pretreatment position, and longer-term factors related to anatomical changes and response to treatment. Images obtained before treatment should be aligned with those taken in the planning CT and they should be a perfect match. Today, many modern Linac devices incorporate daily "cone-beam" CT (CBCT) that uses mega-voltage X-rays for treatment verification, but this imaging is not sufficient to distinguish soft tissue structures. However, since these images are used to adapt treatment plans to patients' daily anatomy and reduce intra-fractional shifts, they are considered suitable for image-guided radiotherapy. Before applying RT daily, CBCT needs to be reviewed before each treatment and two or at least one radiotherapy technique is required for this procedure. When there is an anatomical difference between CBCT and planning CT, radiotherapy technicians should notify the radiation oncologist and medical physicist. At this stage, it is necessary to decide whether to continue treatment with this difference or if a new CBCT is required. Each of these steps delays patient treatment and causes a significant increase in the workload of the RT department. They also lead to the growth of ML in parallel with the training program in radiation oncology. In addition to being able to cope with the growing workload of existing staff, innovations in modern technology and the ability to benefit from it depend on accessibility to sufficient human resources. ML has also been used for replanning; i.e., defining candidate patients for adaptive RT. With machine learning, the machine will shorten the time it takes to train the staff as they can "learn". (47). Classifiers and clustering algorithms have been developed to predict patients that will most benefit from updated plans during fractionated RT, based on anatomical and dosimetric variations (tumor shrinkage, patient weakening, edema, etc.) (48-49). However, it should also be taken into account that since ML learns about previous patients, their plans, and adaptive RT from the available data, it will mimic past protocols rather than determining the ideal time for replanning.

Patient Follow-up

ML also has the potential to change the way radiation oncologists follow up patients undergoing definitive treatments. After surgery, the tumor may disappear during imaging and tumor markers can quickly normalize. In contrast, after RT, the changes in imaging (such as loss of contamination, PET involvement or diffusion restriction, and reduced tumor size) and the response of tumor markers are gradual. These characteristics are monitored regularly over time and response assessment is performed based on the changes that are considered to be indicative of therapeutic efficacy and complemented by clinical experience. This assessment takes time, but if patients not responding to treatment can be predicted earlier, the decision to implement an additional dose of RT or additional systemic treatments can be made earlier, which may improve oncological outcomes. In this context, early studies in the field of radiology are promising. In radiology, quantitative features are extracted to characterize an image based on size and shape, image density, texture, relationships between voxels, and some other characteristics. ML algorithms can be used to associate image-based features with biological observations or clinical results (50-55).

Using techniques for response and survival prediction in radiotherapy patients presents an important opportunity to further improve decision support systems and to provide an objective assessment of the relative benefits of various treatment options for patients. Bryce et al. used ML to evaluate the data from a randomized phase II trial conducted with 95 patients with advanced stage head and neck cancer who underwent RT ± KT and determined the important variables for two-year survival as T stage, N stage, tumor size, tumor resectability, and hemoglobin value. The two-year survival was reported as AUC ROC 0.78 ± 0.05 (56). In another study related to prostate cancer, 119 cases were assessed, the RT dose, dose distribution, and associated biological factors were investigated, and biochemical control and bladder and rectum toxicity were predicted (57).

A promising area for implementing techniques in radiation oncology is to predict toxicity as a valuable decision support system for clinicians (58). The first ML study on toxicity was undertaken in 2009 by Zhang et al. to evaluate the complications of IMRT. A total of 120 plans were created for a head and neck cancer case and a further 256 plans were produced for a prostate cancer case to examine the prediction results of the saliva flow rate and grade 2 rectal bleeding. The absolute error rate in the estimation of the saliva flow rate was calculated as 0.42% and the accuracy of grade 2 rectal bleeding prediction was 97% (59). Pella et al. recorded the clinical and dosimetric data of 321 prostate cancer patients, scored gastrointestinal and genitourinary acute toxicities, and classified the patients according to the mild and severe toxicity categories. Using ML algorithms, the accuracy was determined as AUC 0.7 (60). In another study evaluating rectal toxicity in cervical cancer, the authors predicted grade 2 rectal toxicity based on the dose distribution in the rectum using ML and found AUC as 0.7 (61). In another ML study, they predicted sensorineural hearing loss with 70% accuracy using CT radiomics in cases diagnosed with head and neck cancer and treated with RT + CT (62).

2. Conclusion

In radiation oncology, ML and DL can now be involved in every step from patient consultation to follow-up and contribute to clinicians and society, but there are still many challenges that need to be overcome and many problems to be solved. For ML, large data sets must be created and improved. It is considered that robust models cannot be constructed with data from a single institution, and data sharing is required. In addition, the data collection process should be standardized. Today, the accuracy and quality of data are of great importance since no ML algorithm is able to correct errors in training data.

REFERENCES

1. Meyer P, Noblet V, Mazzara C, et al. Survey on deep learning for radiotherapy. *Comput Biol Med*. 2018;98:126–46.
2. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015;521:436–44.
3. Jarrett D, Stride E, Vallis K, et al. Applications and limitations of machine learning in radiation oncology. *Br J Radiol*. 2019;92:20190001.
4. Boldrini L, Bibault J-E, Masciocchi C, et al. Deep learning: A review for the radiation oncologist. *Front Oncol*. 2019;9:977.
5. Chen C, He M, Zhu Y, et al. Five critical elements to ensure the precision medicine. *Cancer Metastasis Rev*. 2015;34:313–8.
6. Bibault JE, Giraud P, Burgun A. Big data and machine learning in radiation oncology: State of the art and future prospects. *Cancer Lett*. 2016;382:110–7.
7. Lambin P, van Stiphout RG, Starmans MH, et al. Predicting outcomes in radiation oncology—multifactorial decision support systems. *Nat Rev Clin Oncol.* 2013;10:27–40.
8. Quinlan JR. Induction of decision trees. *Mach Learn.* 1986;1:81–106.
9. Langley P, Sage S. Induction of selective Bayesian classifiers. Proceedings of the Tenth International Conference on Uncertainty in Artificial Intelligence; San Francisco, USA. *Morgan Kaufmann Publishers Inc*.; 1994.
10. Patrick EA, Fischer FP. A generalized k-nearest neighbor rule. Inf Control. 1970;16:128–52.
11. Vapnik V. (1982). *Estimation of Dependences Based on Empirical Data*. New York: Springer-Verlag.
12. Rumelhart DE, McClelland J. (1986) *Parallel distributed processing: Explorations in the microstructure of cognition*. Cambridge: MIT Press.
13. Miller AA. Developing an ontology for radiation oncology, master of information and communication technology. *Research thesis, School of Information Systems and Technology, University of Wollongong;* 2012.
14. Feng M, Valdes G, Dixit N, et al. Machine learning in radiation oncology: opportunities, requirements, and needs. *Front Oncol*. 2018;8:110.
15. Valdes G, Luna JM, Eaton E, et al. MediBoost: a patient stratification tool for interpretable decision making in the era of precision medicine. *Sci Rep*. 2016;6:37854.
16. Caruana R, Lou Y, Gehrke J, et al. Intelligible models for healthcare: Predicting pneumonia risk and hospital 30-day readmission. Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; Sidney, Australia. *ACM*; 2015.

17. Caruana R, Niculescu-Mizil A. An empirical comparison of supervised learning algorithms. Proceedings of the 23rd International Conference on Machine Learning; Pittsburgh, Pennsylvania. *ACM*; 2006.
18. Fernández-Delgado M, Cernadas E, Barro S, et al. Do we need hundreds of classifiers to solve real world classification problems? *J Mach Learn Res*. 2014;15:3133–81.
19. Roques TW. Patient selection and radiotherapy volume definition — can we improve the weakest links in the treatment chain? *Clin Oncol*. 2014;26:353–5.
20. Sharp G, Fritscher KD, Pekar V, et al. Vision 20/20: perspectives on automated image segmentation for radiotherapy. Med Phys. 2014;41:050902.
21. Peressutti D, Schipaanboord B, van Soest J, et al. TU-AB-202-10: how effective are current atlas selection methods for atlas-based Auto-Contouring in radiotherapy planning? *Med Phys*. 2016;43:3738–9.
22. Lustberg T, van Soest J, Gooding M, et al. Clinical evaluation of atlas and deep learning based automatic contouring for lung cancer. *Radiother Oncol*. 2018;126:312–7.
23. Ibragimov B, Xing L. Segmentation of organs-at-risks in head and neck CT images using convolutional neural networks. *Med Phys*. 2017;44:547–57.
24. Guo Y, Gao Y, Shen D. Deformable MR prostate segmentation via deep feature learning and sparse patch matching. *IEEE Trans Med Imaging* 2016; 35:1077–89.
25. Kamnitsas K, Ledig C, Newcombe VFJ, et al. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. Med Image Anal. 2017;36:61–78.
26. Men K, Zhang T, Chen X, et al. Fully automatic and robust segmentation of the clinical target volume for radiotherapy of breast cancer using big data and deep learning. *Phys Med*. 2018;50:13–19.
27. Cardenas CE, McCarroll RE, Court LE, et al. Deep learning algorithm for auto-delineation of high-risk oropharyngeal clinical target volumes with built-in dice similarity coefficient parameter optimization function. *Int J Radiat Oncol Biol Phys* 2018;101:468–78.
28. Men K, Dai J, Li Y. Automatic segmentation of the clinical target volume and organs at risk in the planning CT for rectal cancer using deep dilated convolutional neural networks. *Med Phys*. 2017;44:6377–89.
29. Boutilier JJ, Craig T, Sharpe MB, et al. Sample size requirements for knowledge-based treatment planning. Med Phys. 2016;43:1212–21.
30. Schreibmann E, Fox T. Prior-knowledge treatment planning for volumetric arc therapy using feature-based database mining. *J Appl Clin Med Phys*. 2014;15:4596.
31. Tol JP, Delaney AR, Dahele M, et al. Evaluation of a knowledge-based planning solution for head and neck cancer. *Int J Radiat Oncol Biol Phys*. 2015;91:612–20.
32. Shiraishi S, Tan J, Olsen LA, et al. Knowledge-based prediction of plan quality metrics in intracranial stereotactic radiosurgery. *Med Phys*. 2015;42:908–17.
33. Moore KL, Brame RS, Low DA, et al. Experience-based quality control of clinical intensity-modulated radiotherapy planning. *Int J Radiat Oncol Biol Phys*. 2011;81:545–51.
34. Ahmed S, Nelms B, Gintz D, et al. A method for a priori estimation of best feasible DVH for organs-at-risk: validation for head and neck VMAT planning. *Med Phys*. 2017;44:5486–97.
35. Fried DV, Chera BS, Das SK. Assessment of PlanIQ feasibility DVH for head and neck treatment planning. *J Appl Clin Med Phys*. 2017;18:245–50.
36. McIntosh C, Welch M, McNiven A, et al. Fully automated treatment planning for head and neck radiotherapy using a voxel-based dose prediction and dose mimicking method. *Phys Med Biol*. 2017;62:5926–44.
37. Valdes G, Simone CB, Chen J, et al. Clinical decision support of radiotherapy treatment planning: a data-driven machine learning strategy for patient-specific dosimetric decision making. *Radiother Oncol*. 2017;125:392–7.
38. Chanyavanich V, Das SK, Lee WR, et al. Knowledge-based IMRT treatment planning for prostate cancer. *Med Phys*. 2011;38:2515–22.
39. Kusters JMAM, Bzdusek K, Kumar P, et al. Automated IMRT planning in Pinnacle: a study in head-and-neck cancer. *Strahlenther Onkol*. 2017;193:1031–8.
40. Rowbottom CG, Webb S, Oldham M. Beam-orientation customization using an artificial neural network. *Phys Med Biol*. 1999;44:2251.
41. Llacer J, Li S, Agazaryan N, et al. Non-coplanar automatic beam orientation selection in cranial IMRT: a practical methodology. *Phys Med Biol*. 2009;54:1337-68.
42. Valdes G, Morin O, Valenciaga Y, et al. Use of TrueBeam developer mode for imaging QA. *J Appl Clin Med Phys*. 2015;16:322-33.
43. Li Q, Chan MF. Predictive time-series modeling using artificial neural networks for Linac beam symmetry: an empirical study. *Ann N Y Acad Sci*. 2017;1387:84–94.
44. Valdes G, Scheuermann R, Hung CY, et al. A mathematical framework for virtual IMRT QA using machine learning. *Med Phys*. 2016;43:4323–34.
45. Valdes G, Chan MF, Lim SB, et al. IMRT QA using machine learning: a multi-institutional validation. *J Appl Clin Med Phys*. 2017;18:279–84.
46. Carlson JN, Park JM, Park SY, et al. A machine learning approach to the accurate prediction of multi-leaf collimator positional errors. *Phys Med Biol*. 2016;61:2514-31.
47. Boon IS, Yong TPT, Boon CS. Assessing the role of artificial intelligence (AI) in clinical oncology: utility of machine learning in radiotherapy target volume delineation. *Medicines (Basel*); 2018;5:E131.
48. Guidi G, Maffei N, Vecchi C, et al. Expert system classifier for adaptive radiation therapy in prostate cancer. *Australas Phys Eng Sci Med*. 2017;40:337–48.

49. Guidi G, Maffei N, Meduri B, et al. A machine learning tool for re-planning and adaptive RT: a multicenter cohort investigation. *Phys Med*. 2016;32:1659–66.
50. Tseng HH, Luo Y, Cui S, et al. Deep reinforcement learning for automated radiation adaptation in lung cancer. *Med Phys*. 2017;44:6690–705.
51. Varfalvy N, Piron O, Cyr MF, et al. Classification of changes occurring in lung patient during radiotherapy using relative γ analysis and hidden Markov models. *Med Phys*. 2017;44:5043–50.
52. Oakden-Rayner L, Carneiro G, Bessen T, et al. Precision radiology: predicting longevity using feature engineering and deep learning methods in a radiomics framework. *Sci Rep*. 2017;7:1648.
53. Lao J, Chen Y, Li ZC, et al. A deep learning-based radiomics model for prediction of survival in glioblastoma multiforme. *Sci Rep*. 2017;7:10353.
54. Li Z, Wang Y, Yu J, et al. Deep learning based radiomics (DLR) and its usage in noninvasive IDH1 prediction for low grade glioma. *Sci Rep*. 2017;7:5467.
55. Cha KH, Hadjiiski L, Chan HP, et al. Bladder cancer treatment response assessment in CT using radiomics with deep- learning. *Sci Rep*. 2017;7:8738.
56. Bryce TJ, Dewhirst MW, Floyd CE, et al. Artificial neural network model of survival in patients treated with irradiation with and without concurrent chemotherapy for advanced carcinoma of the head and neck. *Int J Radiat Oncol Biol Phys*. 1998;41:339–45.
57. Gulliford SL, Webb S, Rowbottom CG, et al. Use of artificial neural networks to predict biological outcomes for patients receiving radical radiotherapy of the prostate. *Radiother Oncol*. 2004;71:3–12.
58. Kang J, Schwartz R, Flickinger J, et al. Machine learning approaches for predicting radiation therapy outcomes: a clinician's perspective. *Int J Radiat Oncol Biol Phys*. 2015;93:1127–35.
59. Zhang HH, D'Souza WD, Shi L, et al. Modeling plan-related clinical complications using machine learning tools in a multiplan IMRT framework. *Int J Radiat Oncol Biol Phys*. 2009;74:1617–26.
60. Pella A, Cambria R, Riboldi M, et al. Use of machine learning methods for prediction of acute toxicity in organs at risk following prostate radiotherapy. *Med Phys*. 2011;38:2859–67.
61. Zhen X, Chen J, Zhong Z, et al. Deep convolutional neural network with transfer learning for rectum toxicity prediction in cervical cancer radiotherapy: a feasibility study. *Phys Med Biol*. 2017;62:8246–63.
62. Vial A, Stirling D, Field M, et al. The role of deep learning and radiomic feature extraction in cancer-specific predictive modelling: a review. *Transl Cancer Res*. 2018;7:803–16.