SA–CME Information

Applications of Artificial Intelligence in Head-and-Neck Radiation Therapy

Description

The purpose of this article is to review recent advancements in artificial intelligence (AI) as they specifically pertain to head-and-neck radiation oncology. The main focus will be applications and clinical impact of these techniques. Specifically, this article addresses autosegmentation of organs at risk (OARs), autosegmentation of target volumes, treatment planning and predictive dose calculation, image-guided adaptive radiation therapy, prognosis and outcome prediction, and quality assurance.

Learning Objectives

After completing this activity, participants will be able to:

1. Understand the fundamentals of artificial intelligence
2. Explore how artificial intelligence can be applied to head-and-neck cancer radiation therapy in the following venues: autosegmentation of organs-at-risk, autosegmentation of target volumes, treatment planning and predictive dose calculation, image-guided adaptive radiation therapy, prognosis and outcome prediction, and quality assurance
3. Recognize limitations to avoid pitfalls in and maximize the potential of artificial intelligence

Authors

Adam C. Riegel, PhD, DABR, is an associate chief physicist, Department of Radiation Medicine, Northwell Health, Lake Success, NY.

OBTAINING CREDITS

Instructions: To successfully earn credit, participants must complete the activity during the valid credit period.

1. Review this article in its entirety.
2. Visit www.appliedradiology.org/SAM.
3. Login to your account or (new users) create an account.
4. Complete the post test and review the discussion and references.
5. Complete the evaluation.
6. Print your certificate.

Date of release and review: March 1, 2021
Expiration date: February 28, 2023
Estimated time for completion: 1 hour

Disclosures: No authors, faculty, or individuals at the Institute for Advanced Medical Education (IAME) or *Applied Radiation Oncology* who had control over the content of this program have relationships with commercial supporters.

Accreditation/Designation Statement: The IAME is accredited by the Accreditation Council for Continuing Medical Education (ACCME) to provide continuing medical education for physicians. The IAME designates this journal-based activity for a maximum of 1 AMA PRA Category 1 Credit™. Physicians should only claim credit commensurate with the extent of their participation in the activity. These credits qualify as SA-CME credits for ABR diplomates, based on the criteria of the American Board of Radiology.

Commercial Support: None

As part of this CME activity, the reader should reflect on how it will impact his or her personal practice and discuss its content with colleagues.

Applications of Artificial Intelligence in Head and Neck Radiation Therapy

Adam C. Riegel, PhD, DABR

The role of artificial intelligence (AI) in radiation oncology has increased dramatically in the past 5 years, touching nearly every aspect of our field. Artificial intelligence can be broadly defined as "the use of a machine (computer) to perform tasks that typically require human thought."[1] In radiation oncology, these tasks were previously limited to highly repetitive actions that could be scripted in common programming languages. Recently, though, as the widespread accessibility of powerful computing resources has enabled the utilization of machine learning, researchers are finding novel applications for AI in many aspects of radiation oncology previously thought impossible.

Machine learning techniques can be defined as algorithms that yield output from a given input without specific instructions. The algorithms "learn" by detecting underlying patterns in the input data. This period of learning is called model training. Training can be

Dr. Riegel is an associate chief physicist, Department of Radiation Medicine, Northwell Health, Lake Success, NY. Disclosure: The author has no conflicts of interest to disclose and has not received outside funding for the production of this original manuscript. No part of this article has been previously published elsewhere.

supervised (where the model is generated to produce known output) or unsupervised (where the model determines its own output based on the data itself).[1] Deep learning is a specific type of machine learning that utilizes an artificial neural network that models human neurocognitive design to simulate human thought and understanding. Deep-learning architectures have several hidden layers that process input data through deeper levels of abstraction to learn patterns and produce output.[2] The patterns or "features" are often complex and nonlinear in nature.[2,3] One of the more common types of deep-learning methodologies for image-based tasks is a convolutional neural network (CNN). First introduced to nonmedical image classification by Krizhevsky et al in 2012,[4] CNNs convolve input data with multiple filters or "kernels" to produce progressively more abstract representations of the input data. Many AI applications in radiation oncology utilize some variation of the CNN.

The purpose of this article is to review recent advancements in AI as they specifically pertain to head and neck radiation oncology. Although some technical details regarding AI techniques will be discussed, the main focus will be application and clinical impact of these techniques. Specifically, this article will focus on the following

applications: autosegmentation of organs at risk (OARs), autosegmentation of target volumes, treatment planning and predictive dose calculation, image-guided adaptive radiation therapy, prognosis and outcome prediction, and quality assurance. For more granular detail about AI methodologies in radiation oncology, the reader is referred to the excellent review articles cited here.[1-3,5,6]

Organ-at-Risk Segmentation

OAR segmentation is an ideal task for automation due to its repetitive nature and the common geometric properties of normal anatomy shared among all members of the population. Furthermore, manual delineation of head and neck OARs is tedious and prone to variation among multiple observers.[7] Early attempts at automatic OAR segmentation involved a posteriori region-growing and edge-detection approaches. Following this early work, automatic OAR segmentation was accomplished with single- or multi-atlas-based techniques that utilized deformable image registration to warp contours from a similar atlas patient to the current patient.[8,9] Such atlas-based approaches are now widely available as commercial products by multiple vendors.

Recently, researchers have assessed the use of machine learning in OAR segmentation with impressive results. Several authors have shown improvements

in volume overlap with "ground truth" contours using models trained on computed tomography (CT) datasets.[10-16] In these studies, "ground truth" contours typically consisted of expert contours or consensus contours from public databases. Metrics to compare automatically segmented volumes with ground truth included variations of the Dice similarity coefficient index (DSC), Hausdorff distance, or average surface distance.[6] Not only were OARs more accurate, they were also generated faster using AI.

Van Dijk et al reported significantly improved results using a deep-learning approach relative to atlas-based autosegmentation for 19 of 22 head and neck OARs.[11] The deep-learning architecture consisted of multiple CNNs trained on a relatively large database of more than 500 CT image sets. Furthermore, human observers found that CNN-based contours generated fewer obvious errors than atlas-based autosegmentation (9% vs 30%, respectively) and, for most OARs, were found to require less correction than atlas-based segmentation. This algorithm is one of the few commercial deep-learning segmentation tools and is known by the brand name DLCExpert (Mirada Medical, Ltd.).

Though less common than CT, researchers have also investigated autosegmentation on other imaging modalities such as MRI. Yang et al focused on segmentation of the parotid gland in pre- and post-treatment treatment MRI to better quantify changes in parotid volume. The authors used support vector machine classification to train the 15-patient model. T1- and T2-weighted postcontrast MRIs were acquired pretreatment and at 3-, 6-, and 12-month intervals after treatment. Overlap with physician-drawn parotid contours was over 90% for both parotid glands in follow-up MRI scans and autosegmented contours highlighted a 25% reduction in parotid volume at 3 months.[17]

Target Segmentation

Automatic segmentation of gross tumor volumes (GTVs) and clinical target volumes (CTVs) is more difficult than that of OARs due to the inherently abnormal nature of the anatomy, but potentially yields benefits in reducing delineation variability and increasing efficiency. Cardenas et al have written several papers on automatic CTV delineation in head and neck cancers.[18-20] In their first 2018 publication, the authors used manually segmented GTVs from 52 node positive and negative oropharyngeal patients to train a deep-learning model to generate high-risk CTVs with a nonuniform margin. The deep-learning model showed good overlap with manually segmented ground truth CTVs (mean DSC range from 0.755 to 0.840 for all pathologies).[20] In their second paper, the authors used a CNN to train a CTV-generating model on 285 oropharyngeal patients and compared its performance to atlas-based segmentation. Overlap DSC was 0.816 for deep learning and 0.739 for atlas-based segmentation.[19] In their most recent paper, the authors focused specifically on nodal CTVs by training a new model with 51 head and neck patients of varying primary site. Node level volumes were contoured and used as input in the CNN deep-learning architecture. The DSC for nodal CTVs ranged from 0.843 to 0.909 compared to ground truth and, qualitatively, more than 99% were scored as acceptable by a panel of 3 experts.[18]

In a study aimed at contouring GTV (split into primary and nodal volumes) and CTV for nasopharyngeal cancer, Men et al set a deep deconvolutional neural network (with an added deconvolution step at the end of the network to restore some high-resolution features) against a conventional CNN.[21] The authors demonstrated significantly better overlap with ground truth for all targets using the experimental architecture (82.6% and 80.9% vs 73.7% and 72.3%

for CTV and GTV primary, respectively), but nodal GTV lagged in performance at 62.3%. Although this was better than 33.7% with the conventional CNN, the authors highlighted a few reasons for the deficiency, including lack of clear anatomical boundaries, variable target locations, and poor contrast on CT.[21]

The lack of contrast on CT can be mitigated by adding a second modality with supplementary information such as positron emission tomography (PET)/CT. Guo et al used a 3D CNN to develop a model to segment GTV using CT simulation and registered diagnostic PET/CT. The model was trained using 140 patients with squamous cell carcinoma whose PET was deformably registered to simulation CT. Three models were created: CT alone, PET alone, and CT simulation with registered PET. The combined PET/CT model outperformed CT alone and PET alone by 0.4 and 0.05 in mean overlap metrics, respectively, demonstrating the advantage of incorporating functional information into the model.[22] Berthon et al proposed a decision tree that, through machine learning, would select from multiple automatic segmentation algorithms. The decision tree was tested on 20 oropharyngeal patients and segmented GTVs overlapped with manually drawn ground truth with a DSC of 0.77.[23]

Treatment Planning and Predictive Dose Calculation

The first step of the treatment planning process is CT simulation. For years, physicists have been researching ways to replace CT simulation with MRI simulation because of MRI's superior soft-tissue contrast. The largest hurdle in replacing CT with MRI is arguably the loss of electron density information provided in CT that is used in dose calculation. Like automatic segmentation, earlier approaches to "synthetic" CT generation (electron density

maps produced from MRI) began with thresholding methods and atlas-based deformable image registration. Researchers are now using deep learning to produce synthetic CTs. Dinkla et al rationalize the need for MRI-based planning in head and neck cancer in the context of MR-guided linear accelerators. Thirty-four head and neck cancer patients received CT simulation and large field of view T2-weighted MRI on a 3T scanner. Average absolute errors were 75 ± 9, 214 ± 26, 35 ± 3, and 130 ± 24 HU for body, bone, soft tissue, and air, respectively. The authors speculate that HU values for bone in synthetic CT are slightly lower than actual CTs due to registration errors between CT and MRI. Dose distributions calculated on synthetic CT were within 1% of dose calculated on actual CT voxel by voxel.[24] Klages et al performed a similar study comparing 23 patients in two general adversarial networks to generate synthetic CT. Mean absolute HU errors and dosimetric errors were comparable to Dinkla et al, but the authors found that combining results from three orthogonal views decreased HU errors.[25]

Automated treatment planning is another potential application of AI. Currently, commercial knowledge-based planning systems use the relative geometry of targets and OARs in previously treated patients to predict the dose-volume histograms for de novo patients.[26] Machine learning is being applied to dose prediction as well, with the goal that more accurate dose prediction will yield inverse optimization parameters to speed up the planning process. Chen et al used 70 nasopharyngeal cancer patients who had been treated with 6 MV step-and-shoot intensity-modulated radiation therapy (IMRT) to train a CNN-based dose prediction model. A unique aspect of this model was that the authors tested a general model against a modification that specifically identified out-of-field voxels to potentially increase accuracy near the edges of beams. The

model was then tested against 10 additional nasopharyngeal patients and the predicted dose was compared voxel by voxel against the clinical treatment plan. For most regions of interest, the models performed comparably, but the "out-of-field" modification significantly improved agreement for the smaller regions of interest such as chiasm, lenses, and optic nerves.[27] In a more recent paper, the same group compared two CNN-based dose prediction models specifically for helical tomotherapy. The models were CResDevNet and a standard U-Net architecture. Using 136 nasopharyngeal treatment plans and 24 validation plans, the models were tested against 60 patients. The mean absolute error with clinical plans was between $3.2 \pm 2.5\%$ and $3.7 \pm 2.9\%$ for the CResDevNet and U-Net, respectively. CResDevNet also had a slight advantage with the majority of OARs when the overlap of dose-volume histogram curves was measured.[28]

Adaptive Radiation Therapy

Adaptive radiation therapy is a specialized form of image-guided radiation therapy frequently used in head and neck treatment sites due to the significant anatomical changes that can occur over the course of treatment. Currently, adaptive radiation therapy usually uses an "offline" approach. The physician may set a predefined trigger point, perhaps halfway through treatment, where the plan will be re-evaluated based on localization imaging such as cone-beam CT (CBCT) and adapted to current anatomy if necessary. The physician may also call for ad hoc adaptation based on changes seen in image guidance or on-treatment visits. In offline adaptive therapy, the patient receives a new CT simulation and a new plan is generated for the remaining fractions. This is extremely time-consuming and labor-intensive as the entire treatment planning process must be repeated with the new CT. Although training AI models requires substantial time upfront,

increased operational speed yields significant benefit to offline adaptive radiation therapy and opens the possibility of "online" adaptive radiation therapy where the plan is adapted immediately after localization imaging is acquired and the patient remains on the table. Without the increased computational speed that machine learning provides, the feasibility of online adaptation is questionable.

Tong et al investigated the use of adversarial networks for OAR segmentation on both CT simulation and low-field MRI acquired on an MR-guided linear accelerator for online adaptation of image-guided therapy. The CT model was trained on 48 patients from the RTOG 522 dataset and the MRI model was trained on 25 MRI volumes acquired on the MRIdian system (ViewRay). The authors found that the adversarial network that included multiple integrated neural networks (SC-GAN-DenseNet) performed better than other models for both CT and low-field MRI. This is particularly notable given the low signal-to-noise environment of low-field MRI and the short contouring time for the deep-learning model (approximately 14 seconds compared to 30 minutes for the comparable model-based algorithm).[16] Although not currently a focus in head and neck applications, intrafraction motion management may also benefit from fast contour propagation in MR-guided linear accelerators with continuous image monitoring during treatment.[29]

Guidi et al deformably registered daily localization MVCT or kV CBCT to CT simulations to measure the changes in dose due to changing anatomy over time. The authors focused on parotid glands as they are prone to substantial changes during the course of treatment. Using a support vector machine, changes in parotid volume were classified into categories ranging from "Correct Treatment" where planning was not necessary to "Suggested Replanning" where changes indicate

replanning would mitigate suboptimal dosimetric changes. The authors included additional classifications to alert users of abnormal changes in volume data that are either extreme anatomical fluctuations or artifactual. The authors found that, by the fourth week of treatment, approximately 55% of patients required retreatment. The authors validated their approach by comparing their classification to physician judgment and found good concordance between the model and physicians.[30]

One of the few commercial online image-guided adaptive therapy solutions is the Ethos platform by Varian. Built on Halcyon hardware, the Ethos system contains proprietary deep-learning AI that drives online adaptation of the plan from iteratively reconstructed kV CBCT. In a recent publication, the authors (Varian employees) vaguely describe the deep-learning algorithm as CNN-based similar to U-net and DenseNet. After targets and organs at risk are propagated to current anatomy, the plan is reoptimized while the patient is on the treatment table.[31] Though numerous abstracts were presented at recent national meetings about this adaptive platform, few manuscripts on clinical user experiences have yet been published.

Outcome Prediction

Outcome prediction is another application of AI in head and neck radiation therapy. Several authors have explored normal tissue complication probability (NTCP) prediction of xerostomia and dysphagia with machine learning methods.[32-34] Lee et al used quality of life surveys to identify the most influential predictive factors in a multivariable xerostomia model in squamous cell carcinoma and nasopharyngeal carcinoma. Interestingly, the authors found that, in addition to ipsilateral and contralateral parotid dose, features such as age, T-stage, financial status, and education were also significant predictors of xerostomia.[32,33] Dean et al built three

dysphagia machine learning models based on prior work[35] with 173 patients for training and 90 patients for validation from a variety of head and neck disease sites and several institutions. Dysphagia was scored using CTCAE version 3 and dose to the pharyngeal mucosa was considered along with other clinical factors. The authors compared models trained with dose-volume data alone vs inclusion of spatial dose information. The authors found that spatial dose information did not improve NTCP modeling of dysphagia and therefore recommended the standard model that includes dose-volume information only, with the caveat that different spatial dose metrics may produce different results.[34]

Research in radiomics, the study of hundreds or thousands of subtle features within regions of interest contoured on diagnostic imaging, has been accelerated by machine learning algorithms.[36] Ren et al developed a model to differentiate between stage I-II and stage III-IV squamous cell carcinoma by extracting 970 radiomic features from multisequence MRI.[37] Van Dijk et al used the 90th percentile of the MRI signal from pretreatment T1-weighted MRI to predict xerostomia.[38] Gabryś et al compared conventional NTCP prediction models of xerostomia with machine learning models that included a variety of user-selected radiomic and dosiomic features. Xerostomia was split into early, late, and long-term time periods, with acceptable predictive success occurring only for long-term toxicity. Small parotids with steep dose gradients in the lateral direction were more prone to xerostomia. The authors suggested this may be caused by the changes in anatomy during treatment, pulling the smaller glands close to high dose regions.[39] If this is true, such a finding would support the need for adaptive therapy as described above.

Quality Assurance

The use of machine learning in medical physics quality assurance procedures is

rooted in the idea that physics resources are scarce and should be allocated where they can make the most impact. In other words, tasks that can be automated should be automated so that physicists, like physicians, can concentrate on tasks that truly require expert human judgment. Several authors have investigated the utility of machine learning in identifying plan parameter outliers,[40] finding erroneous contours,[41] calculating output factors in proton therapy,[42] predicting MLC leaf position errors,[43] and predicting gamma index passing rates in IMRT QA.[44-48]

In their 2016 paper, Valdes et al state their intent to create virtual IMRT QA[46] where planners would be able to predict the gamma indices of a given plan before running the QA, potentially avoiding overmodulated plans and ultimately saving time. This is particularly applicable to head and neck treatment plans as they tend to be more complex than other anatomical sites. The authors began by training a generalized linear model with Poisson regression and LASSO regularization on nearly 500 Eclipse-based treatment plans (Varian) for a variety of sites.[46] The authors then augmented this model with portal dosimetry measurements from a different institution[47] and, in their most recent publication, updated their 500-plan model using a CNN called VGG-16. The authors found comparable results between the CNN and their Poisson-regression model, although the CNN yielded several advantages over the Poisson model including calculation speed (after model training) and independence from user-selected features.[48]

Limitations of Artificial Intelligence in Radiation Therapy

There is tremendous potential in AI-based approaches to solving our most pressing problems in head and neck radiation therapy. There are, however, limitations to what AI can currently

accomplish. Image reconstruction parameters, for example, significantly impact the performance of image-based AI algorithms for CT,[49,50] PET,[51] and MRI.[52] Quality of autosegmentation and radiomics analysis is dependent on the accuracy of ground truth contours, which are typically drawn by physicians prone to intra- and interobserver variation.[53] Deep-learning algorithms are abstract by nature and the source of errant results can be difficult to pinpoint.[2,3,54] Large high-quality datasets are required for adequate training and validation of deep-learning algorithms to prevent overfitting.[2,5,54] Even with an adequate sample, deep-learning techniques can be fooled by subtle changes in imaging,[2] which is potentially troubling if imaging artifacts occur. Given the "black box" nature of artificial intelligence, thorough validation procedures are required to ensure models are yielding reasonable results. Regulators are understandably cautious about certifying such powerful and complex tools for clinical use, which may explain the relatively limited number of commercially available AI tools.

Conclusion

Advancements in AI continue at a rapid pace. Given the plethora of digital data generated for patients undergoing head and neck radiation therapy, radiation oncology is well positioned to harness the power of machine learning and deep learning to improve decision-support algorithms, autosegmentation, treatment planning, outcome prediction, and quality assurance. Although few commercial products exist using AI technology, it is only a matter of time until such products are available. It will be incumbent upon us as medical professionals to familiarize ourselves with the basics of AI so we may shine a light in the "black box" and provide the most intelligent care (artificial or otherwise) to our patients.

REFERENCES

1. Deig CR, Kanwar A, Thompson RF. Artificial intelligence in radiation oncology. *Hematol Oncol Clin North Am.* 2019;33(6):1095-1104. doi:10.1016/j.hoc.2019.08.003

2. Meyer P, Noblet V, Mazzara C, Lallement A. Survey on deep learning for radiotherapy. *Comput Biol Med.* 2018;98:126-146. doi:10.1016/j.compbiomed.2018.05.018

3. Cardenas CE, Yang J, Anderson BM, Court LE, Brock KB. Advances in auto-Segmentation. *Semin Radiat Oncol.* 2019;29(3):185-197. doi:10.1016/j.semradonc.2019.02.001

4. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Adv Neural Inf Process Sys.* 2012;60(6):1097-1105.

5. Wang C, Zhu X, Hong JC, Zheng D. Artificial intelligence in radiotherapy treatment planning: present and future. *Technol Cancer Res Treat.* 2019;18:1-11. doi:10.1177/1533033819873922

6. Vrtovec T, Močnik D, Strojan P, Pernuš F, Ibragimov B. Auto-segmentation of organs at risk for head and neck radiotherapy planning: from atlas-based to deep learning methods. *Med Phys.* 2020;47(9):e929-e950. doi:10.1002/mp.14320

7. Riegel AC, Berson AM, Destian S, et al. Variability of gross tumor volume delineation in head and neck cancer using CT and PET/CT fusion. *Int J Radiat Oncol Biol Phys.* 2006;65(3):726-732. doi:10.1016/j.ijrobp.2006.01.014

8. Lim JY, Leech M. Use of auto-segmentation in the delineation of target volumes and organs at risk in head and neck. *Acta Oncol (Madr).* 2016;55(7):799-806. doi:10.3109/0284186X.2016.1173723

9. Kosmin M, Ledsam J, Romera-Paredes B, et al. Rapid advances in auto-segmentation of organs at risk and target volumes in head and neck cancer. *Radiother Oncol.* 2019;135:130-140. doi:10.1016/j.radonc.2019.03.004

10. Liang S, Tang F, Huang X, et al. Deep-learning-based detection and segmentation of organs at risk in nasopharyngeal carcinoma computed tomographic images for radiotherapy planning. *Eur Radiol.* 2019;29(4):1961-1967. doi:10.1007/s00330-018-5748-9

11. van Dijk LV, Van den Bosch L, Aljabar P, et al. Improving automatic delineation for head and neck organs at risk by deep learning contouring. *Radiother Oncol.* 2020;142:115-123. doi:10.1016/j.radonc.2019.09.022

12. Ibragimov B, Xing L. Segmentation of organs-at-risks in head and neck CT images using convolutional neural networks: *Med Phys.* 2017;44(2):547-557. doi:10.1002/mp.12045

13. Zhu W, Huang Y, Zeng L, et al. Anatomy-Net: Deep learning for fast and fully automated whole-volume segmentation of head and neck anatomy. *Med Phys.* 2019;46(2):576-589. doi:10.1002/mp.13300

14. Chan JW, Kearney V, Haaf S, et al. A convolutional neural network algorithm for automatic segmentation of head and neck organs at risk using deep lifelong learning. *Med Phys.* 2019;46(5):2204-2213. doi:10.1002/mp.13495

15. Tong N, Gou S, Yang S, Ruan D, Sheng K. Fully automatic multi-organ segmentation for head and neck cancer radiotherapy using shape representation model constrained fully convolutional neural networks. *Med Phys.* 2018;45(10):4558-4567. doi:10.1002/mp.13147

16. Tong N, Gou S, Yang S, Cao M, Sheng K. Shape constrained fully convolutional DenseNet with adversarial training for multiorgan segmentation on head and neck CT and low-field MR images. *Med Phys.* 2019;46(6):2669-2682. doi:10.1002/mp.13553

17. Yang X, Wu N, Cheng G, et al. Automated segmentation of the parotid gland based on atlas registration and machine learning: A longitudinal mri study in head and neck radiation therapy. *Int J Radiat Oncol Biol Phys.* 2014;90(5):1225-1233. doi:10.1016/j.ijrobp.2014.08.350

18. Cardenas CE, Beadle BM, Garden AS, et al. Generating high-quality lymph node clinical target volumes for head and neck cancer radiotherapy using a fully automated deep learning-based approach. *Int J Radiat Oncol.* 2021;109(3):801-812. doi:10.1016/j.ijrobp.2020.10.005

19. Cardenas CE, Anderson BM, Aristophanous M, et al. Auto-delineation of oropharyngeal clinical target volumes using 3D convolutional neural networks. *Phys Med Biol.* 2018;63(21). doi:10.1088/1361-6560/aae8a9

20. 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(2):468-478. doi:10.1016/j.ijrobp.2018.01.114

21. Men K, Chen X, Zhang Y, et al. Deep deconvolutional neural network for target segmentation of nasopharyngeal cancer in planning computed tomography images. *Front Oncol.* 2017;7(DEC):1-9. doi:10.3389/fonc.2017.00315

22. Guo Z, Guo N, Gong K, Zhong S, Li Q. Gross tumor volume segmentation for head and neck cancer radiotherapy using deep dense multi-modality network. *Phys Med Biol.* 2019;64(20). doi:10.1088/1361-6560/ab440d

23. Berthon B, Evans M, Marshall C, et al. Head and neck target delineation using a novel PET automatic segmentation algorithm. *Radiother Oncol.* 2017;122(2):242-247. doi:10.1016/j.radonc.2016.12.008

24. Dinkla AM, Florkow MC, Maspero M, et al. Dosimetric evaluation of synthetic CT for head and neck radiotherapy generated by a patch-based three-dimensional convolutional neural network. *Med Phys.* 2019;46(9):4095-4104. doi:10.1002/mp.13663

25. Klages P, Benslimane I, Riyahi S, et al. Patch-based generative adversarial neural network models for head and neck MR-only planning. *Med Phys.* 2020;47(2):626-642. doi:10.1002/mp.13927

26. Tol JP, Delaney AR, Dahele M, Slotman BJ, Verbakel WFAR. Evaluation of a knowledge-based planning solution for head and neck cancer. *Int J Radiat Oncol Biol Phys.* 2015;91(3):612-620. doi:10.1016/j.ijrobp.2014.11.014

27. Chen X, Men K, Li Y, Yi J, Dai J. A feasibility study on an automated method to generate patient-specific dose distributions for radiotherapy using deep learning. *Med Phys.* 2019;46(1):56-64. doi:10.1002/mp.13262

SA-CME (see page 6)

28. Liu Z, Chen X, Men K, Yi J, Dai J. A deep learning model to predict dose–volume histograms of organs at risk in radiotherapy treatment plans. *Med Phys*. Published online 2020. doi:10.1002/mp.14394

29. Eppenhof KAJ, Maspero M, Savenije MHF, et al. Fast contour propagation for MR-guided prostate radiotherapy using convolutional neural networks. *Med Phys*. 2020;47(3):1238-1248. doi:10.1002/mp.13994

30. Guidi G, Maffei N, Meduri B, et al. A machine learning tool for re-planning and adaptive RT: A multicenter cohort investigation. *Phys Medica*. 2016;32(12):1659-1666. doi:10.1016/j.ejmp.2016.10.005

31. Archambault Y, Boylan C, Bullock D, et al. Making on-line adaptive radiotherapy possible using artificial intelligence and machine learning for efficient daily re-planning. *Med Phys Int J*. 2020;8(2):77-86.

32. Lee TF, Liou MH, Huang YJ, et al. LASSO NTCP predictors for the incidence of xerostomia in patients with head and neck squamous cell carcinoma and nasopharyngeal carcinoma. *Sci Rep*. 2014;4:1-8. doi:10.1038/srep06217

33. Lee TF, Chao PJ, Ting HM, et al. Using multivariate regression model with least absolute shrinkage and selection operator (LASSO) to predict the incidence of xerostomia after intensity-modulated radiotherapy for head and neck cancer. *PLoS One*. 2014;9(2). doi:10.1371/journal.pone.0089700

34. Dean J, Wong K, Gay H, et al. Incorporating spatial dose metrics in machine learning-based normal tissue complication probability (NTCP) models of severe acute dysphagia resulting from head and neck radiotherapy. *Clin Transl Radiat Oncol*. 2018;8:27-39. doi:10.1016/j.ctro.2017.11.009

35. Dean JA, Wong KH, Welsh LC, et al. Normal tissue complication probability (NTCP) modelling using spatial dose metrics and machine learning methods for severe acute oral mucositis resulting from head and neck radiotherapy. *Radiother Oncol*. 2016;120(1):21-27. doi:10.1016/j.radonc.2016.05.015

36. Giraud P, Giraud P, Gasnier A, et al. Radiomics and machine learning for radiotherapy in head and neck cancers. *Front Oncol*. 2019;9(MAR):1-13. doi:10.3389/fonc.2019.00174

37. Ren J, Tian J, Yuan Y, et al. Magnetic resonance imaging based radiomics signature for the preoperative discrimination of stage I-II and III-IV head and neck squamous cell carcinoma. *Eur J Radiol*. 2018;106(March):1-6. doi:10.1016/j.ejrad.2018.07.002

38. van Dijk LV, Thor M, Steenbakkers RJHM, et al. Parotid gland fat related Magnetic Resonance image biomarkers improve prediction of late radiation-induced xerostomia. *Radiother Oncol*. 2018;128(3):459-466. doi:10.1016/ j.radonc.2018.06.012

39. Gabryś HS, Buettner F, Sterzing F, Hauswald H, Bangert M. Design and selection of machine learning methods using radiomics and dosiomics for normal tissue complication probability modeling of xerostomia. *Front Oncol*. 2018;8:1-20. doi:10.3389/fonc.2018.00035

40. Azmandian F, Kaeli D, Dy JG, et al. Towards the development of an error checker for radiotherapy treatment plans: a preliminary study. *Phys Med Biol*. 2007;52(21):6511-6524. doi:10.1088/0031-9155/52/21/012

41. McIntosh C, Svistoun I, Purdie TG. Groupwise conditional random forests for automatic shape classification and contour quality assessment in radiotherapy planning. *IEEE Trans Med Imaging*. 2013;32(6):1043-1057. doi:10.1109/TMI.2013.2251421

42. Sun B, Lam D, Yang D, et al. A machine learning approach to the accurate prediction of monitor units for a compact proton machine. *Med Phys*. 2018;45(5):2243-2251. doi:10.1002/mp.12842

43. Carlson JNK, Park JM, Park SY, Park JI, Choi Y, Ye SJ. A machine learning approach to the accurate prediction of multi-leaf collimator positional errors. *Phys Med Biol*. 2016;61(6):2514-2531. doi:10.1088/0031-9155/61/6/2514

44. Wootton LS, Nyflot MJ, Chaovalitwongse WA, Ford E. Error detection in intensity-modulated radiation therapy quality assurance using radiomic analysis of gamma distributions. *Int J Radiat Oncol Biol Phys*. 2018;102(1):219-228. doi:10.1016/j.ijrobp.2018.05.033

45. Nyflot MJ, Thammasorn P, Wootton LS, Ford EC, Chaovalitwongse WA. Deep learning for patient-specific quality assurance: Identifying errors in radiotherapy delivery by radiomic analysis of gamma images with convolutional neural networks. *Med Phys*. 2019;46(2):456-464. doi:10.1002/mp.13338

46. Valdes G, Scheuermann R, Hung CY, Olszanski A, Bellerive M, Solberg TD. A mathematical framework for virtual IMRT QA using machine learning. *Med Phys*. 2016;43(7):4323-4334. doi:10.1118/1.4953835

47. Valdes G, Chan MF, Lim SB, Scheuermann R, Deasy JO, Solberg TD. IMRT QA using machine learning: a multi-institutional validation. *J Appl Clin Med Phys*. 2017;18(5):279-284. doi:10.1002/acm2.12161

48. Interian Y, Rideout V, Kearney VP, et al. Deep nets vs expert designed features in medical physics: an IMRT QA case study. *Med Phys*. 2018;45(6):2672-2680. doi:10.1002/mp.12890

49. Shafiq-Ul-Hassan M, Zhang GG, Latifi K, et al. Intrinsic dependencies of CT radiomic features on voxel size and number of gray levels. *Med Phys*. 2017;44(3):1050-1062. doi:10.1002/mp.12123

50. Zhao W, Zhang W, Sun Y, et al. Convolution kernel and iterative reconstruction affect the diagnostic performance of radiomics and deep learning in lung adenocarcinoma pathological subtypes. *Thorac Cancer*. 2019;10(10):1893-1903. doi:10.1111/1759-7714.13161

51. Galavis PE, Hollensen C, Jallow N, Paliwal B, Jeraj R. Variability of textural features in FDG PET images due to different acquisition modes and reconstruction parameters. *Acta Oncol (Madr)*. 2010;49(7):1012-1016. doi:10.3109/0284186X.2010.498437

52. Mayerhoefer ME, Szomolanyi P, Jirak D, Materka A, Trattnig S. Effects of MRI acquisition parameter variations and protocol heterogeneity on the results of texture analysis and pattern discrimination: An application-oriented study. *Med Phys*. 2009;36(4):1236-1243. doi:10.1118/1.3081408

53. Pavic M, Bogowicz M, Würms X, et al. Influence of inter-observer delineation variability on radiomics stability in different tumor sites. *Acta Oncol (Madr)*. 2018;57(8):1070-1074. doi:10.1080/0284186X.2018.1445283

54. Thompson RF, Valdes G, Fuller CD, et al. Artificial intelligence in radiation oncology imaging. *Int J Radiat Oncol Biol Phys*. 2018;102(4):1159-1161. doi:10.1016/j.ijrobp.2018.05.070