

# Artificial Intelligence-guided Prediction of Dental Doses prior Toplanning of Head and Neck Radiotherapy: Technical Development and Pilot Study of Feasibility

**Jason W. Chan**

University of California San Francisco

**Nicole Hohenstein**

University of California San Francisco

**Colin Carpenter**

Siris Medical Inc.

**Adam J. Pattison**

Siris Medical Inc.

**Olivier Morin**

University of California San Francisco

**Gilmer Valdes**

University of California San Francisco

**Maria Thompson**

University of California San Francisco

**Jennifer Perkins**

University of California San Francisco

**Timothy D. Solberg**

Food and Drug Administration

**Sue S. Yom** (✉ [sue.yom@ucsf.edu](mailto:sue.yom@ucsf.edu))

University of California San Francisco <https://orcid.org/0000-0002-0779-7476>

---

## Research

**Keywords:** Radiotherapy dosage, tooth extraction, head and neck cancer, computer-assisted radiotherapy, radiation dosimetry, machine learning, artificial intelligence, osteoradionecrosis

**Posted Date:** July 1st, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-38929/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

## ***Background:***

The aim of this research was to develop a novel artificial intelligence (AI)-guided clinical decision support (CDS) system, to predict radiation doses to subsites of the mandible using diagnostic CT scans acquired before planning of head and neck radiotherapy (RT).

## ***Methods:***

A dose classifier was trained using RT plans from 86 oropharyngeal cancer patients; the test set consisted of an additional 20 plans. The classifier was trained to predict whether mandible subsites would receive a mean dose >50 Gy. The AI predictions were prospectively evaluated and compared to those of a specialist head and neck radiation oncologist for 9 patients. Positive predictive value (PPV), negative predictive value (NPV), Pearson's correlation coefficient, and Lin's concordance correlation coefficient were calculated to compare the AI predictions to those of the physician.

## ***Results:***

In the test dataset, the AI predictions had a PPV of 0.95 and NPV of 0.88. For 9 patients evaluated prospectively, there was a strong correlation between the predictions of the AI algorithm and physician ( $p = 0.72$ ,  $p < 0.001$ ). Comparing the AI algorithm versus the physician, the PPVs were 0.82 versus 0.25, and the NPVs were 0.94 versus 1.0, respectively. Concordance between physician estimates and final planned doses was 0.62; this was 0.71 between AI-based estimates and final planned doses.

## ***Conclusions:***

An AI-guided CDS tool to predict dental dosimetry prior to head and neck RT was built, validated, and prospectively tested. AI-guided decision support increased precision and accuracy of dental dose estimates and improved the quality of pre-RT dental assessment.

# Background

Dental decay and osteoradionecrosis (ORN) are major long-term sequelae of head and neck radiotherapy (HN RT) but there is little concrete basis to counsel patients in advance about their individual risk of potential complications [1–3]. Dentists and oral surgeons often base their pre-RT assessments on the prescription doses that are anticipated by the consulting radiation oncologist but are rarely provided with actual dosimetry for specific subsites of the mandible. It is generally advised that in areas which receive a high dose of radiation (e.g. > 50 Gy), there should be a lifelong relative contraindication against dentoalveolar surgery including procedures such as dental extractions, dental implants, periodontal surgery, and surgical endodontic therapy [1]. Thus prophylactic dental extractions are frequently performed before RT begins for any teeth that are felt to be compromised and which lie within a region anticipated to receive a high dose of radiation, or teeth that are assessed to be non-restorable by any

means. Prophylactic dental extractions require weeks to heal and can delay RT treatment by several weeks; the removal of teeth can also compromise patients' speech and diet and potentially worsen their nutritional status just prior to beginning RT.

The utilization of data to drive improved RT treatment planning is a rapidly maturing field. In one application, a technique known as knowledge-based planning (KBP) uses historical planning information to automate RT treatment planning of new patients' cases [4]. Machine-learning algorithms such as logistic regression, linear models with penalty regularization, decision trees, random forest, gradient boosting, and deep learning approaches have helped advance KBP [5, 6]. However, KBP approaches require full-scale contouring of clinical cases and are not able to provide anticipatory clinical decision support (CDS) prior to initiation of contouring and radiotherapy planning of a specific case [4]. The alternative CDS framework is intended to augment expert opinion prior to the point of contouring and/or planning and has been shown to help decision making at the point of care [7], while avoiding some of the known limitations of AI including lack of clinical context or unintended bias [8].

In this study, we developed an AI-based CDS tool to predict which sub-regions of the mandible would receive  $D_{\text{mean}} > 50$  Gy, prior to initiation of any actual RT planning activities. This CDS was based on a previously validated machine-learning algorithm which is capable of deriving dosimetry for new oropharyngeal cancer patients based on historical data mining [9]. A process was developed to predict doses to the mandible that would result from the eventual RT treatment plan based on minimal contouring of a diagnostic CT. This accuracy of this algorithm was compared to estimates provided by radiation oncologists specialized in the treatment of head and neck cancer. The CDS-based dose estimation was also evaluated prospectively on 9 patients, compared against the actual dosimetry that resulted from radiotherapy planning.

## Methods

### Clinical workflow of predicting dental dosimetry using diagnostic imaging

The clinical workflow is shown in Fig. 1. Diagnostic computed tomography (CT) scans were imported from the Picture and Archiving System (PACS, Agfa Healthcare, Mortsel, Belgium) into a commercial software (MiM Maestro, MiM Software, Cleveland, Ohio) for manual delineation. In keeping with this institution's usual workflow, the primary and nodal gross tumor volumes (GTV) were co-delineated with neuroradiology input [10]. Per consensus contouring guidelines, a 5 mm expansion was added to the GTV to create the clinical target volume (CTV) and an additional 3 mm margin was added to create the planning tumor volume (PTV); an example is shown in Fig. 3 [11]. The diagnostic CT and associated contours (structure sets) were then exported to the ML system.

# Machine learning model building, patient selection, and definition of mandible subsites on diagnostic CTs

A dataset of 86 previously delivered RT treatment plans, including planning CT datasets with their associated structure sets and treatment plans with 3D dose distributions, was used for training the ML system. A test dataset, independent from the training dataset, was created from 20 patients whose cases were chronologically subsequent to the training dataset. The test set included not only planning CTs, structure sets, and the RT treatment plans, but also associated diagnostic CTs for which the mandible had been contoured. All of these patients were treated with curative intent, using intensity-modulated radiotherapy (IMRT) prescribed to a total dose of at least 60 Gy. The six subsites of the mandible were manually contoured and defined as follows: left posterior (teeth 17–19), left middle (teeth 20–22), left anterior (teeth 23–24), right anterior (teeth 25–26), right middle (teeth 27–29), and right posterior (teeth 30–32).

## Dose prediction to mandible subsites

A custom, autocontouring algorithm used the training dataset to segment mandible subsites in the diagnostic CT scans (Fig. 2). Autocontouring accuracy was assessed using the dosimetric method proposed by Lim et al. on the test dataset [12]; this method assesses contouring performance by comparing the final delivered dose volume (in units of Gray (Gy)) of the manually created structure to the dose-volume of the autocontoured structure. A machine learning (ML) dose prediction model was trained using a 5-fold cross-validation method; performance was subsequently measured on the test dataset. The dose prediction methodology has previously been described in detail [9]. Briefly, the model uses an ensemble method incorporating boosting, which forms a strong classifier through iterative learning of weak classifiers from features obtained from CT images, PTVs, and organs at risk (OARs) [13]. Included features in this model incorporate prior assessment of DICOM images, medical record data (e.g. ICD-9/10 diagnosis codes), radiation-transport parameters (e.g. beam energy), and the treatment intent (e.g. prescription dose). Of particular importance for this project were the geometrical features of the mandible substructures and the organs in close proximity to the mandible (e.g. oral cavity).

The goal for this specific project was to predict the mean dose ( $D_{\text{mean}}$ ) to each of the six mandible subsites and the overall maximum point dose ( $D_{\text{max}}$ , defined as the dose to a point of 0.03 cc) to any part of the mandible. For the mandible subsites, the key dosimetric metric was  $D_{\text{mean}} > 50$  Gy, but  $D_{\text{mean}} > 40$  Gy and  $D_{\text{mean}} > 30$  Gy were also evaluated. Additionally, the dosimetric indices for the GTV, CTV, PTV, and OARs were analyzed and aggregated to form the treatment plan matrix.

The AI-predicted dosimetric indices were compared to the actual mandible dosimetry obtained after RT treatment planning. Positive Predictive Value (PPV) and Negative Predictive Value (NPV) were used to evaluate the performance of the AI-based CDS in predicting  $D_{\text{mean}}$  of  $> 50$  Gy,  $> 40$  Gy, and  $> 30$  Gy to the six subsites of the mandible.

## Prospective, correlative clinical comparison

Head and neck specialist radiation oncologists provided estimates of 9 patients' mandible subsite doses to the dental oncology team, as is current standard practice. These values were compared to the CDS using Pearson's correlation and Lim's concordance analysis.

After the physicians had estimated the mandible doses and the contoured diagnostic CT scans had been sent to the CDS, the patients underwent a CT-based simulation and had standard RT treatment planning. Clinically accepted, deliverable RT treatment plans were finalized using a commercial treatment planning software (Raystation, Raysearch Laboratories, Stockholm, Sweden). The dose to the mandible subsites that was clinically planned to be actually delivered to the patient was compared to the physician-estimated dose and to the dose as predicted by the CDS.

From these data, the PPV and NPV were used to evaluate the ability to predict whether each mandible subsite received  $D_{mean} > 50$  Gy as follows:

		Dmean actually delivered to mandible subsites		
		Condition Positive	Condition Negative	
<b>Dmean estimated by either physician or CDS to mandibular subsites</b>	Test outcome positive	True positive (TP) Total subsites correctly predicted and actually received $D_{mean} > 50$ Gy	False positive (FP) Total subsites with over-estimated dose prediction (did not receive $D_{mean} > 50$ Gy)	<b>Positive Predictive Value (PPV)</b>  $TP / (TP + FP)$
	Test outcome negative	False negative (FN) Total subsites with under-estimated dose prediction (received $D_{mean} > 50$ Gy)	True negative (TN) Total subsites correctly predicted and actually did not receive $D_{mean} > 50$ Gy	<b>Negative Predictive Value (NPV)</b>  $TN / (FN + TN)$

The research was conducted under Institutional Review Board-approved quantitative imaging analysis, #18-25441.

## Results

### Accuracy of dose prediction on mandible subsites for the test set

Performance metrics were determined for predicting  $D_{mean} > 50$  Gy to the autocontoured mandible subsites on diagnostic CT scans. The PPV of estimating  $D_{mean} > 50$  Gy was 0.95 and the NPV was 0.88. To evaluate the dosimetric error propagated by autocontouring, dose prediction with autocontouring was compared to manual contouring on the test set. These results demonstrated relatively small mean

absolute errors (MAEs) in the prediction of mean doses to subsites of the mandible and the maximum dose to the entire mandible. In comparing results of autocontouring to manual contouring, the resulting average MAE was 1.5 Gy and all MAEs were within 3.30 Gy (Table 1).

Table 1  
Mean absolute error (MAE) of dose prediction (in units of Gray (Gy)) in the test set with the use of autocontouring rather than manual contouring of the mandible and its subsites.

<b>Mandible Subsite</b>	<b>Mean absolute error (MAE) (Gy)</b>
Mandible, Anterior Left Mean	0.96
Mandible, Anterior Right Mean	1.05
Mandible, Middle Left Mean	0.75
Mandible, Middle Right Mean	1.38
Mandible, Posterior Left Mean	3.30
Mandible, Posterior Right Mean	2.93
Mandible, Maximum	0.11

## Prospective clinical comparison of CDS and standard of care

A strong correlation was observed between the AI-based CDS- and physician-generated predictions (Pearson’s correlation coefficient  $r = 0.72$ ,  $p < 0.001$ ). Using the delivered treatment plan as the ground truth, the PPVs for the CDS versus the physician were 0.82 versus 0.25, respectively, for predicting  $D_{\text{mean}} > 50$  Gy to the mandible subsites. The NPVs were 0.94 versus 1.0, respectively. Additional comparisons of PPVs and NPVs for several other dose thresholds ( $D_{\text{mean}} > 40$  Gy,  $D_{\text{mean}} > 30$  Gy) are shown in Table 2; the PPV of the CDS predictions were equivalent or superior to the physicians’ estimates at every dose level.

Table 2

Negative predictive values (NPVs) and positive predictive values (PPVs) for the mean doses (D<sub>mean</sub>) delivered to mandible subsites as estimated by the radiation oncology physician versus the AI-based clinical decision support system (CDS); PPVs were equivalent or higher for the CDS at all dose levels. Estimates were for 9 patients whose mandible doses were predicted from contouring of their diagnostic computed tomography scan prior to initiation of any radiotherapy planning.

Dose prediction	Radiation oncology physician		AI-based clinical decision support system	
	NPV	PPV	NPV	PPV
D <sub>mean</sub> > 50 Gy	1.0	0.25	0.94	0.82
D <sub>mean</sub> > 40 Gy	0.97	0.62	0.89	0.95
D <sub>mean</sub> > 30 Gy	0.57	0.85	0.50	0.85

Comparing accuracy, the accuracy of the CDS was not statistically similar to the physician estimates (Wilcoxon rank test,  $p = 0.42$ ). While both the CDS and the physician were accurate for the anterior and posterior submandible structures ( $p < 0.01$ ), when comparing the middle teeth, the CDS was statistically similar to the final delivered dose ( $p < 0.01$ ), whereas the physician estimates were not ( $p = 0.4$ ). Concordance between the physician’s estimates and the delivered treatment plan’s doses was 0.62 and concordance between the CDS and the delivered treatment plan’s doses was 0.71.

## Discussion

In this study, we established a novel application improving the accuracy and efficiency of pre-treatment dental evaluation in patients scheduled for head and neck RT, with the ultimate goal of real-world clinical implementation. The end result of this study was an efficient artificial-intelligence (AI) clinical decision support (CDS) system that accurately predicts radiation doses to six subsites of the mandible. There were multiple phases of this development process. First, we established that dose prediction accuracy could be maintained even if only an extremely limited feature set were contoured on a diagnostic CT scan and used for dose prediction; the only necessary contours were the gross tumor (auto-expanded to planning volumes) and the (autocontoured) mandible. Second, based on these minimally contoured diagnostic CT scans, we evaluated the feasibility of dose prediction using the CDS to predict doses to six subsites of the mandible. Finally, we prospectively evaluated the ability of the CDS to predict whether specific subsites of the mandible would receive a mean dose > 50 Gy.

The CDS produced a similar NPV as compared to a physician’s estimate (0.94 and 1.0, respectively), but had a superior PPV of 0.82 as compared to 0.25. Hence, the CDS was much less likely than the physician to overestimate high doses of radiation to the mandible, typically encountered in the posterior aspects of the mandible. Thus, by increasing the accuracy of the RT dose estimates to the mandible, the CDS could prevent unneeded extraction procedures, contributing to patients’ retention of posterior dentition and better nutritional status. Furthermore, the CDS appeared to produce improved accuracy that was superior to the physicians’ estimates in the middle segments of the mandible, where lower radiation doses are

typically delivered; improved accuracy to these lower-dose regions would allow for more nuanced and accurate pre-radiotherapy evaluation of teeth at an intermediate level of risk.

Qualified dental evaluation is an important preparatory step for all patients who are planning to undergo head and neck radiotherapy. The decision to perform extractions prior to radiotherapy is complex and not entirely dependent on dental dosimetry [14], but in general, the expectation of a high radiation dose to a compromised tooth will prompt a recommendation for prophylactic extraction. Extractions in the posterior mandible where high doses often occur usually involve molar teeth which are frequently deep-rooted or impacted; removal procedures are painful and sometimes risk lengthy or permanent non-healing and/or nerve damage. Following extractions, the initiation of radiotherapy may be delayed by a few to several weeks to allow for healing; this comes at the expense of increasing the total treatment package time and potentially compromising tumor control [15, 16]. By improving the specificity of dose estimation to the mandible and reducing the time spent in the pre-RT dental evaluation process, our AI-based CDS has the potential to reduce these delays in initiating RT and increase the overall efficiency of treatment.

The incidence of ORN after head and neck radiation treatment has varied widely in the literature from 0.4% to up to 56% [2]. Tsai et al. reviewed 402 oropharyngeal cancer patients and found that the incidence of ORN was 7.5% and the median time to develop ORN was 8 months [17]. Additionally, after adjustment for dental status (dentate versus post-extraction), patients who developed ORN had larger volumes of the mandible exposed to high doses of radiation (> 50 Gy) as compared to patients who did not develop ORN. Thus, the threshold of 50 Gy appears to be an important factor in the development of subsequent ORN and serves as a common clinical endpoint driving the decision to enact pre-RT prophylactic extractions. The CDS developed and tested in this project proved to be more accurate in identifying this high-risk situation than the usual standard-of-care physician's estimate and it would allow for improved counseling of patients about future risks based on individualized dosimetry.

One potential limitation of this CDS is the reliance on low-dimensional features that are delineated based on patient geometry. However, from an efficiency standpoint, it was necessary to establish a process that could be implemented in the short window of time prior to actual RT treatment planning and which would not overburden physicians with excessive additional contouring responsibilities. We were able to establish that relying only on an extremely limited contour set consisting of the auto-contoured mandible and the primary tumor delineated by the physician, the compromise in terms of dose-prediction accuracy was relatively small. In this proposed workflow, the only manual contouring truly required on the diagnostic CT scan would be the gross disease, which is auto-expanded to the PTV, while the mandible subsites can be reliably auto-contoured [6, 9]; this would likely be acceptable and feasible in the real-world setting of a busy clinical practice.

## Conclusions

Using an artificial intelligence-based clinical decision support system, doses to high-risk OARs can be predicted using only diagnostic imaging, prior to initiation of actual radiation treatment planning.

Incorporating AI-guided dose prediction would facilitate more precise estimations of risks to specific OARs.

## Declarations

*Ethics approval and consent to participate:* The research was conducted under Institutional Review Board-approved quantitative imaging analysis, #18-25441.

*Consent for publication:* Not applicable.

*Availability of data and materials:* The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

*Competing interests:* CC and AJP are employed by Siris Medical Inc. JWC, NH, OM, GV, MT, JP, TDS, SSY declare no competing interests.

*Funding:* This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

*Authors' contributions:* JC, SY conceived of the initial design and provided overall supervision for the research conduct. JC, NH, CC, AP, and SY collected, analyzed, and interpreted data for the project. OM, GV, MT, JP, TS contributed specialized knowledge to the design of the research and assisted in analysis and interpretation of the data. All authors contributed to writing the manuscript, take personal responsibility for the content of the manuscript, and approved the final manuscript.

*Acknowledgements:* Not applicable.

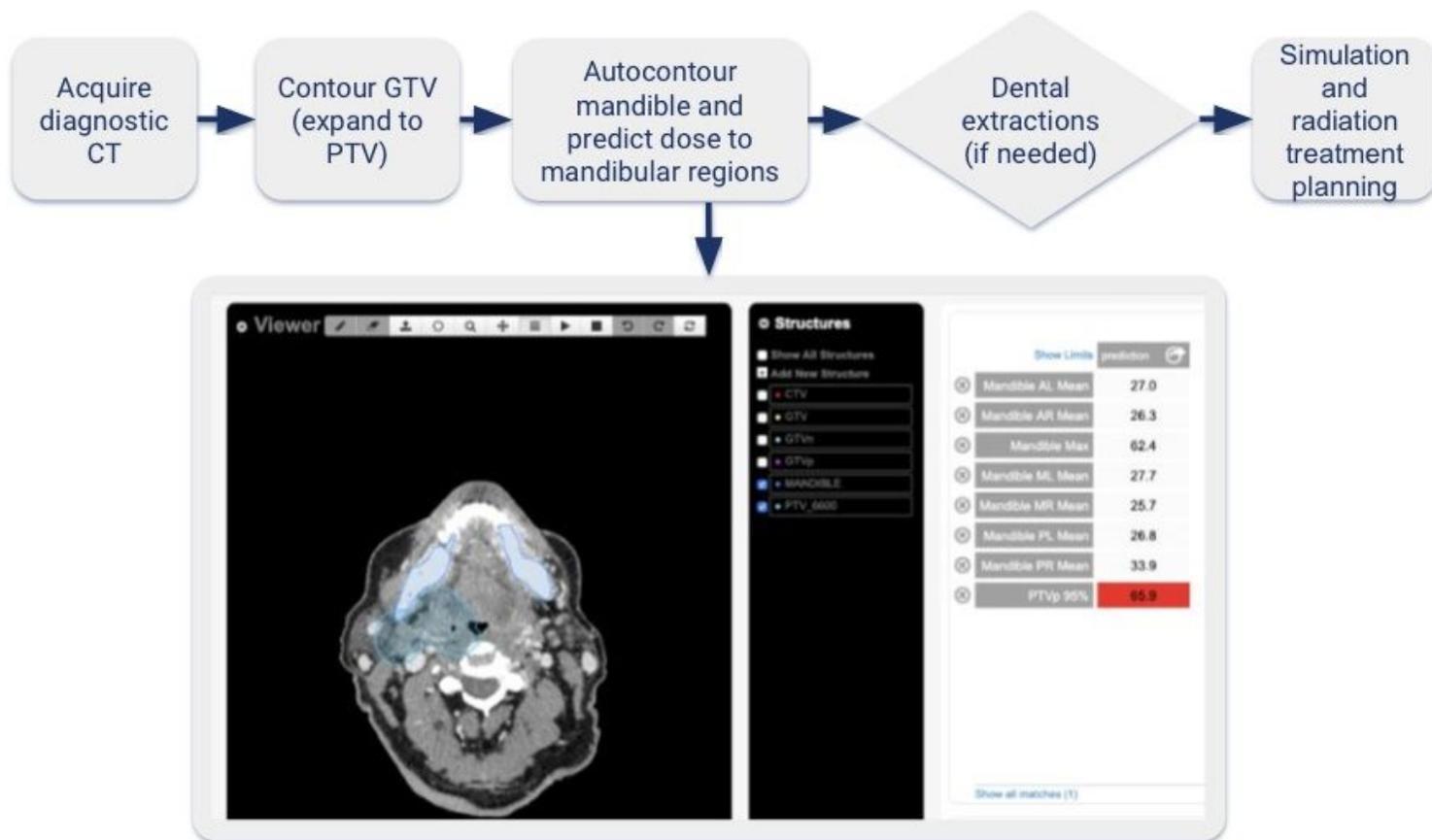
*Authors' information:* Provided above.

## References

1. Devi S, Singh N. Dental care during and after radiotherapy in head and neck cancer. *Natl J Maxillofac Surg.* 2014;5:117–25.
2. Jereczek-Fossa BA, Orecchia R. Radiotherapy-induced mandibular bone complications. *Cancer Treat Rev.* 2002;28:65–74.
3. Koga DH, Salvajoli JV, Alves FA. Dental extractions and radiotherapy in head and neck oncology: review of the literature. *Oral Dis.* 2008;14:40–4.
4. Appenzoller LM, Michalski JM, Thorstad WL, Mutic S, Moore KL. Predicting dose-volume histograms for organs-at-risk in IMRT planning. *Med Phys.* 2012;39:7446–61.
5. Kearney V, Chan JW, Haaf S, Descovich M, Solberg TD. DoseNet: a volumetric dose prediction algorithm using 3D fully-convolutional neural networks. *Phys Med Biol.* 2018;63:235022.

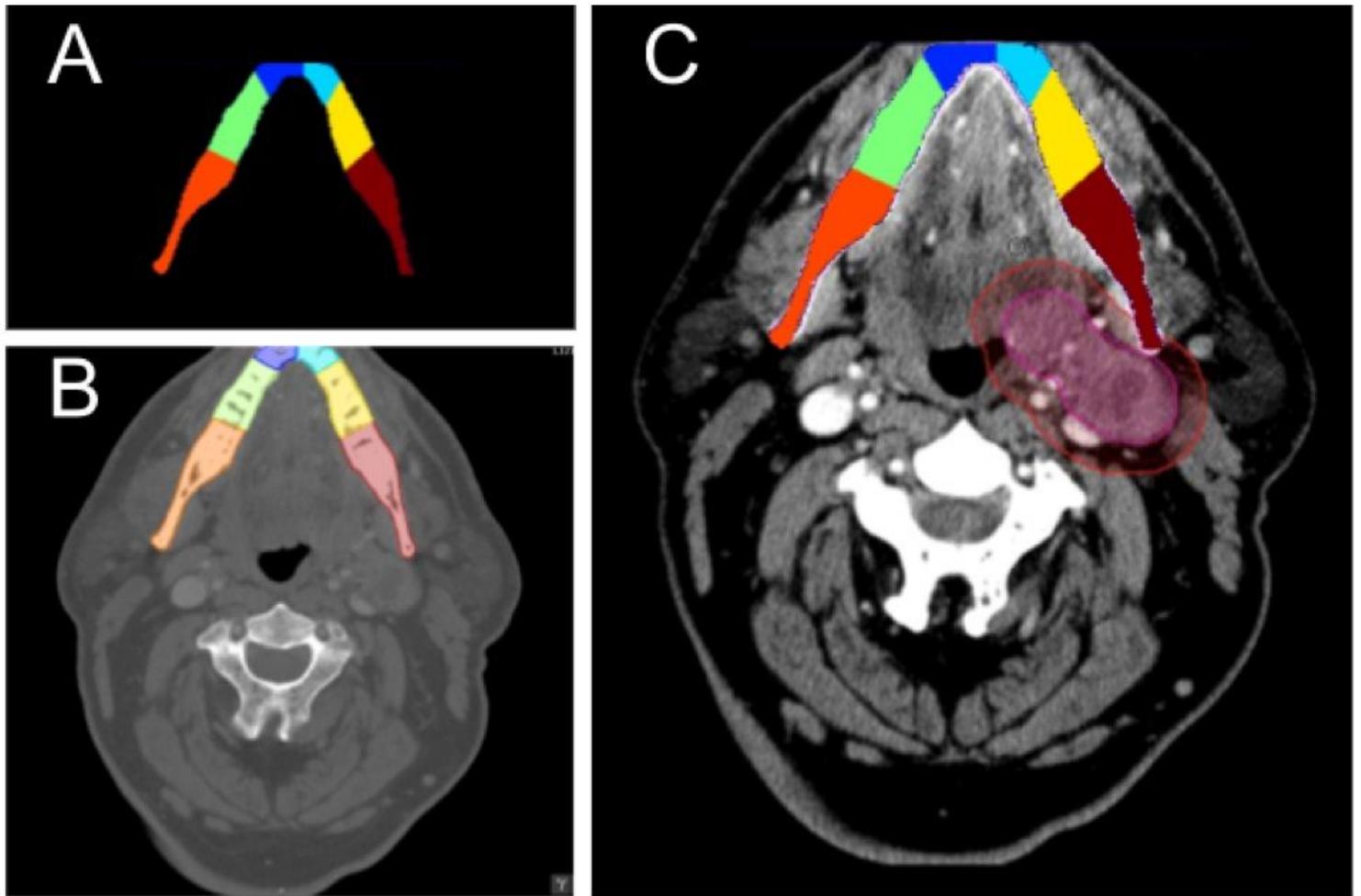
6. Nguyen D, Jia X, Sher D, Lin MH, Iqbal Z, Liu H, et al. 3D radiotherapy dose prediction on head and neck cancer patients with a hierarchically densely connected U-net deep learning architecture. *Phys Med Biol*. 2019;64:065020.
7. Walsh S, de Jong EEC, van Timmeren JE, Ibrahim A, Compter I, Peerlings J, et al. Decision Support Systems in Oncology. *JCO Clin Cancer Inform*. 2019;3:1–9.
8. Cabitza F, Rasoini R, Gensini GF. Unintended Consequences of Machine Learning in Medicine. *JAMA*. 2017;318:517–8.
9. Valdes G, Simone CB 2nd, Chen J, Lin A, Yom SS, Pattison AJ, 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.
10. Braunstein S, Glastonbury CM, Chen J, Quivey JM, Yom SS. Impact of Neuroradiology-Based Peer Review on Head and Neck Radiotherapy Target Delineation. *AJNR Am J Neuroradiol*. 2017;38:146–53.
11. Gregoire V, Evans M, Le QT, Bourhis J, Budach V, Chen A, et al. Delineation of the primary tumour Clinical Target Volumes (CTV-P) in laryngeal, hypopharyngeal, oropharyngeal and oral cavity squamous cell carcinoma: AIRO, CACA, DAHANCA, EORTC, GEORCC, GORTEC, HKNPCSG, HNCIG, IAG-KHT, LPRHHT, NCIC CTG, NCRI, NRG Oncology, PHNS, SBRT, SOMERA, SRO, SSHNO, TROG consensus guidelines. *Radiother Oncol*. 2018;126:3–24.
12. Lim TY, Gillespie E, Murphy J, Moore KL. Clinically Oriented Contour Evaluation Using Dosimetric Indices Generated From Automated Knowledge-Based Planning. *Int J Radiat Oncol Biol Phys*. 2019;103:1251–60.
13. Friedman JH. Greedy function approximation: A gradient boosting machine. *Ann Stat*. 2001;29:1189–232.
14. Adeyemo WL, Taiwo OA, Oderinu OH, Adeyemi MF, Ladeinde AL, Ogunlewe MO. Oral health-related quality of life following non-surgical (routine) tooth extraction: A pilot study. *Contemp Clin Dent*. 2012;3:427–32.
15. Rosenthal DI, Liu L, Lee JH, Vapiwala N, Chalian AA, Weinstein GS, et al. Importance of the treatment package time in surgery and postoperative radiation therapy for squamous carcinoma of the head and neck. *Head Neck*. 2002;24:115–26.
16. Ang KK, Trotti A, Brown BW, Garden AS, Foote RL, Morrison WH, et al. Randomized trial addressing risk features and time factors of surgery plus radiotherapy in advanced head-and-neck cancer. *Int J Radiat Oncol Biol Phys*. 2001;51:571–8.
17. Tsai CJ, Hofstede TM, Sturgis EM, Garden AS, Lindberg ME, Wei Q, et al. Osteoradionecrosis and radiation dose to the mandible in patients with oropharyngeal cancer. *Int J Radiat Oncol Biol Phys*. 2013;85:415–20.

## Figures



**Figure 1**

Clinical workflow of predicting dental dosimetry using diagnostic CTs and an artificial-intelligence-based clinical decision support system. The contoured structures are listed in the middle box and the predicted doses to different regions of the mandible (A=anterior, P=posterior, R=right, L=left) are listed in the leftward box. This example shows a red highlighted box which appears in the dose prediction results when a pre-set quality parameter (in this case, 95% coverage of the planning target volume by prescribed dose) is not met.



**Figure 2**

Example of minimal required contouring on diagnostic computed tomography (CT) scan for dose prediction: (A) autocontoured mandible subsites; (B) mandible subsites overlaid on the diagnostic CT; (C) gross tumor volume (GTV, dark red inner contour) and planning tumor volume (PTV, lighter red outer contour) used for dose prediction as contoured on the diagnostic CT scan.