HHS Public Access

Author manuscript

Int J Radiat Oncol Biol Phys. Author manuscript; available in PMC 2024 August 01.

Published in final edited form as:

Int J Radiat Oncol Biol Phys. 2024 August 01; 119(5): 1569–1578. doi:10.1016/j.ijrobp.2024.02.021.

Cluster-Based Toxicity Estimation of Osteoradionecrosis Via Unsupervised Machine Learning: Moving Beyond Single Dose-Parameter Normal Tissue Complication Probability by Using Whole Dose-Volume Histograms for Cohort Risk Stratification

Seyedmohammadhossein Hosseinian, PhD[*], **Mehdi Hemmati, PhD**[†], **Cem Dede, MD, MSc**[‡], **Travis C. Salzillo, PhD**[‡], **Lisanne V. van Dijk, PhD**[§], **Abdallah S.R. Mohamed, MD, PhD**[‡,‖], **Stephen Y. Lai, MD, PhD**[¶], **Andrew J. Schaefer, PhD**[#], **Clifton D. Fuller, MD, PhD**[‡,#], **on behalf of the Rice/MD Anderson Center for Operations Research in Cancer (CORC) and MD Anderson Head and Neck Cancer Symptom Working Group**

[*]Department of Mechanical and Materials Engineering, University of Cincinnati, Cincinnati, Ohio;

[†]School of Industrial and Systems Engineering, University of Oklahoma, Norman, Oklahoma;

[‡]Department of Radiation Oncology, The University of Texas MD Anderson Cancer Center, Houston, Texas;

[§]Department of Radiation Oncology, University Medical Center Groningen, University of Groningen, Groningen, Netherlands;

[‖]Department of Radiation Oncology, Baylor College of Medicine, Houston, Texas;

[¶]Department of Head and Neck Surgery, The University of Texas MD Anderson Cancer Center, Houston, Texas;

[#]Department of Computational Applied Mathematics & Operations Research, Rice University, Houston, Texas

Abstract

Purpose: Given the limitations of extant models for normal tissue complication probability estimation for osteoradionecrosis (ORN) of the mandible, the purpose of this study was to enrich statistical inference by exploiting structural properties of data and provide a clinically reliable

Corresponding authors: Seyedmohammadhossein Hosseinian, PhD; and Clifton D. Fuller, MD, PhD; s.hosseinian@uc.edu, cdfuller@mdanderson.org.

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ijrobp.2024.02.021.

CRediT (Contributor Roles Taxonomy) Statement

S.H.: Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Software, Supervision, Visualization, Writing – original draft; **M.H.**: Conceptualization, Investigation, Methodology, Software, Visualization, Writing - review & editing; **C.D.**: Conceptualization, Data curation, Investigation, Methodology, Project administration, Writing - review & editing; **T.C.S.**: Data curation, Writing - review & editing; **L.V.v.D.**: Data curation, Investigation, Project administration, Writing - review & editing; **A.S.R.M.**: Conceptualization, Data curation, Investigation, Methodology, Writing - review & editing; **S.Y.L.**: Conceptualization, Resources, Supervision, Writing - review & editing; **A.J.S.**: Funding acquisition, Investigation, Methodology, Resources, Supervision, Writing - review & editing; **C.D.F.**: Conceptualization, Funding acquisition, Investigation, Methodology, Resources, Supervision, Writing - review & editing.

model for ORN risk evaluation through an unsupervised-learning analysis that incorporates the whole radiation dose distribution on the mandible.

Methods and Materials: The analysis was conducted on retrospective data of 1259 patients with head and neck cancer treated at The University of Texas MD Anderson Cancer Center between 2005 and 2015. During a minimum 12-month post-therapy follow-up period, 173 patients in this cohort (13.7%) developed ORN (grades I to IV). The (structural) clusters of mandibular dose-volume histograms (DVHs) for these patients were identified using the K-means clustering method. A soft-margin support vector machine was used to determine the cluster borders and partition the dose-volume space. The risk of ORN for each dose-volume region was calculated based on incidence rates and other clinical risk factors.

Results: The K-means clustering method identified 6 clusters among the DVHs. Based on the first 5 clusters, the dose-volume space was partitioned by the soft-margin support vector machine into distinct regions with different risk indices. The sixth cluster entirely overlapped with the others; the region of this cluster was determined by its envelopes. For each region, the ORN incidence rate per preradiation dental extraction status (a statistically significant, nondose related risk factor for ORN) was reported as the corresponding risk index.

Conclusions: This study presents an unsupervised-learning analysis of a large-scale data set to evaluate the risk of mandibular ORN among patients with head and neck cancer. The results provide a visual risk-assessment tool for ORN (based on the whole DVH and preradiation dental extraction status) as well as a range of constraints for dose optimization under different risk levels.

Introduction

Osteoradionecrosis (ORN) of the mandible is a debilitating side effect of radiation therapy for patients with head and neck cancer (HNC).[1–5] ORN is commonly defined as non-healing bone for a period of at least 3 months due to exposure to radiation.[6–8] Despite its low prevalence,[9–11] ORN may severely affect the quality of life of surviving HNC patients; it is often accompanied by pain, dysesthesia, dysgeusia, and difficulties in mastication and speech, and in case of progression, it may lead to infection, bone fracture, and intraoral or extraoral fistulae.[1,6,12–14] The importance of implementing ORN prevention strategies in modern practice, to enhance the overall quality of life for HNC survivors, underscores the need for developing reliable risk assessment models for ORN. These models can serve as valuable tools to inform and guide dose optimization/adaptation in radiation therapy planning in the quest for more favorable treatment outcomes and patient well-being.

Normal tissue complication probability (NTCP) is a widely accepted prediction tool for radiation toxicity.[15–18] NTCP models aim to provide the likelihood of treatment-induced complications based on the radiation dose delivered to organs at risk (OARs) as well as other clinical risk factors. The current paradigm of NTCP modeling for HNCs relies on supervised-learning (classification) methods, most popularly the classical logistic regression method.[19] During the past decade, a vast number of classification-based NTCP models have been proposed for various HNC treatment-induced complications.[11,20–28] The classification methods, however, have a major limitation for NTCP modeling—due to multicollinearity of dose features—which undermines clinical reliability of these models. Distribution of

radiation dose over an OAR is commonly represented by a dose-volume histogram (DVH). As an OAR's DVH depicts a distribution, the dose-related features for these models (ie, DVH values) are highly correlated. As a result, the entire DVH curve is summarized into a single number (eg, mean dose), which completely disregards the distribution of radiation dose and the fact that clinically different DVH curves—with different toxicity outcomes—may share a feature value. Single-parameter evaluation of dose distributions has been reported as a caveat of early NTCP models (eg, LKB) for clinical use,[16] yet it has persisted throughout the subsequent data-driven models, due to the multicollinearity issue.

In addition, the best performance of classification methods concerns separable data.[29] In general, these methods aim to learn the best separator of 2 training classes from observed clinical data; they perform well if the data are at least softly separable (ie, with relatively few violations). The effect of data inseparability on classification accuracy is shown in Appendix A (Supplementary Materials). Data inseparability is a common characteristic of most NTCP models in radiation therapy. Due to the complexity of side effect development mechanisms, training classes (ie, safe radiation plans vs those leading to complications) are not generally expected to be separable by DVH features. From available data of mandibular ORN development among patients with HNC,[30] it can be observed that the 2 classes of ORN and non-ORN patients are highly inseparable by radiation dose to the mandible; see Appendix B (Supplementary Materials). The inseparability observation is aligned with an evidence-based pathogenesis theory of ORN, suggesting that radiation arteritis leads to the development of a hypocellular, hypovascular, and hypoxic environment, and consequently ORN.[6,7] This implies that a decisive factor in ORN development is the radiation dose received by mandibular arteries, including arterioles and capillaries. Capturing the dose delivered to such a fine network is impractical with the current technologies. On the other hand, DVHs summarize the spatially (3-dimensional [3D]) distributed radiation dose into 1-dimensional data. For statistical inference, such a "summary" is a sufficient statistic only if radiosensitivity of the OAR is spatially homogeneous. However, homogeneous distribution of arteries across the mandible is an unrealistic assumption. This suggests that, while mandibular DVHs are the best available clinical indicator of ORN development, they generally do not contain sufficient information to separate the 2 classes (ORN vs non-ORN) rigorously as intended in the advanced classification methods.

The limitations of supervised-learning methods motivate alternative analyses to build more informative and clinically reliable models for evaluation of the risk of ORN. This article presents a novel unsupervised-learning approach for side effect risk assessment and the results of the analysis on retrospective data of a large cohort of HNC patients for ORN.

Methods and Materials

Data

The analysis was conducted on a data set compiled and made publicly available by van Dijk et al.[30] The data are composed of anonymized (retrospective) clinical records and the mandibular DVHs of 1259 HNC patients treated by radiation therapy alone or in combination with surgery and/or chemotherapy at The University of Texas MD Anderson Cancer Center between 2005 and 2015, after an institutional review board approval

(RCR030800). The clinical records include age, sex, cancer sub-site, T-stage, N-stage, preradiation (within 6 weeks before treatment) dental extraction status, smoking status and pack-years, chemotherapy treatment, definitive versus postoperative radiation therapy, and the mandible bone volume. The reported DVH values include D2%, D5% to D95% in 5% increments, D97%, D98%, and D99% as well as V5 Gy to V70 Gy in 5 Gy increments; the minimum, maximum, and mean dose (Gy) delivered to the mandible are also included. Out of the 1259 patients, 1086 (86.3%) have been reported ORN-free during the minimum of 12 months of posttherapy follow-up; 173 (13.7%) patients are reported to have developed ORN (grades I to IV) during the follow-up period. The data are publicly available at https://doi.org/10.6084/m9.figshare.13568207.

The study cohort constituted part of a larger initiative called the “big data RT HNC.” Exclusions from the study criteria consisted of individuals with previously documented head and neck irradiation, a history of salivary gland cancer, and patients with a survival or follow-up duration of less than 1 year. The prescribed radiation dose for primary tumors ranged from 68 to 72 Gy for definitive treatment (typically administered as 2.12 Gy in 33 fractions over 5 days per week), 60 to 66 Gy for postoperative indications (typically administered as 2 Gy in 30–33 fractions over 5 days per week), and 57 Gy for elective lymph node levels (typically administered as 1.72 Gy in 33 fractions). The standard radiation protocol involved a 6 MV radiation source and a nominal dose rate of 600 Monitor-Units per minute. Among the 1259 patients in this cohort, 123 patients (10%) were treated with 3-dimensional Conformal Radiation Therapy , 891 patients (71%) with IMRT, 224 patients (18%) with VMAT, and 21 patients (2%) with IMPT. The median follow-up duration for all patients in the study was 57 months, ranging from 12 to 174 months. Most patients had oropharyngeal cancer, accounting for 66% of cases, followed by oral cavity cancer (15%) and laryngeal cancer (13%). A significant majority of the patients were male (83%). The patients were treated by different physicians. A complete description of the data has been presented in the study by van Dijk et al.[11]

Analysis

This article presents a secondary analysis of the data set described previously; van Dijk et al[11] have used this data to train a multivariate logistic regression NTCP model for mandibular ORN (any grade). Their univariate analysis, followed by Akaike Information Criterion-based forward stepwise selection, has identified D30% and preradiation dental extraction (PDE) status as the model features; PDE is a nominal (binary) feature indicating no/edentulous versus dental extractions. Although their statistical analysis rigorously identifies radiation intensity and dental extraction as significant predictive factors of ORN development, due to multicollinearity of the dose features, the DVH curve is reduced to D30% in their model. We present a cluster-based analysis that provides more clinical insight about the risk of ORN development as a function of radiation distribution (ie, the whole DVH) and dental extraction status.

The cluster-based analysis takes advantage of the structural properties of mandibular DVHs to divide them into disjoint subgroups (of similar radiation plans), each spanning a distinct region of the dose-volume space. It then renders a risk index for each region based on

the observed incidence rate of ORN (depending on PDE status) in the associated DVH subgroup. The analysis was conducted in 3 steps. Steps 1 and 2 were performed using the scikit-learn library of Python.[31]

Step 1: The K-means clustering method[29] was used to identify the inherent clusters of the mandibular DVHs. To this end, each DVH curve was represented as an array of dose values (ie, Dx% for x% ranging from 2% to 99%); the inertia (within-cluster sum-of-squares) criterion was used to identify an optimal number of clusters.

Step 2: To partition the dose-volume space with respect to these clusters, the shared borders of adjacent clusters were identified by a soft-margin support vector machine (SVM).[29] The analysis was conducted pointwise (for each Dx% individually) to obtain smooth borders among the clusters with minimal violations.

Step 3: For each region of the dose-volume space, spanned by an identified cluster, the ORN incidence rate within the cluster, per PDE status was reported as a point estimate of the associated probability of ORN (risk index) along with the corresponding 95% confidence interval. As a result, each dose-volume region was associated with 2 risk indices (and confidence intervals), for no/edentulous and dental extractions, separately.

The implementation of the method (code) and input data are publicly available at https://github.com/s-hosseinian/Radiation_Therapy_NTCP_ORN.

Results

Figure 1 illustrates the inertia plot of the K-means clustering method. Based on this figure, K = 5 leads to the smallest meaningful within-cluster sum-of-squares. However, the sixth cluster identified by the clustering method (through K = 6) is structurally different from the first 5 clusters and provides additional insight. Therefore, we considered both cases.

Figure 2 illustrates the identified clusters by the K-means clustering method, color-coded for better visualization. A distinction between the results for K = 5 and K = 6 concerns the sixth cluster, depicted in black in Figure 2b. This cluster is mainly composed of a group of DVHs with high dose (above 60 Gy) imposed over most of the mandible volume.

As the next step of the analysis, a soft-margin SVM was used to identify the cluster borders and partition the dose-volume space; SVM renders the best separator of 2 classes. As the analysis was performed pointwise, the best separating point (with minimum violations) was identified for every pair of adjacent clusters for each Dx%. With K = 5, because the black cluster was considered a part of the red cluster and highly overlapping with the others in the most left and right parts of the plot, SVM failed to identify its border in the [D2%, D5%] and [D90%, D99%] intervals. However, with K = 6 and exclusion of the sixth (black) cluster, SVM was able to identify the cluster borders and partition the dose-volume space completely. For the sixth cluster, we determined the corresponding dose-volume region by its envelopes (ie, pointwise minimum and maximum dose). Figure 3 illustrates the cluster borders by dashed black lines.

For each part of the dose-volume space, characterized by the borderlines of the identified clusters, 2 risk indices (and 95% confidence intervals) were calculated based on the observed incidence rate of ORN in the corresponding cluster per PDE status. Tables 1 and 2 present the results for K = 5 and K = 6, respectively. In these tables, PDE = 0 indicates no/edentulous dental extractions, and PDE = 1 concerns performed dental extractions. We have labeled the first 5 clusters according to their order from the lower left to the upper right corners of the DVH plots; cluster No. 6 is the overlapping cluster, illustrated in black in Figures 3b and 3c. The columns “ORN incidence” indicate the number of patients with a positive ORN record out of the total number of patients in that cluster with the same PDE status, and “risk index (95% CI)” reports the corresponding ratio (ie, point estimate of the probability of ORN) as well as the associated 95% confidence interval. Figure 4 shows the risk indices, denoted by r, on the dose-volume regions identified in the previous step for K = 6, as it provided a more accurate partition of the dose-volume space. Each region spans the area between 2 consecutive cluster borders, except for the region of cluster No. 6, which was determined by its envelopes. The borders of the sixth cluster are depicted by dotted lines in this figure. The coordinates of the borderlines are provided in Appendix C (Supplementary Materials).

Model validation

Cross validation was performed to examine accuracy and generalizability of the model. The data were randomly divided into a training set (70%, 880 patients) and a test set (30%, 379 patients). To ensure that the split ratio was preserved in PDE = 0 and PDE = 1 categories, the training set included 70% of the data in each category individually. In the training phase, DVH clusters of the training set and associated risk indices (as well as 95% confidence intervals) were identified as explained before for both K = 5 and K = 6. The K-means clustering method of scikit-learn library outputs the cluster centroids.[31] For each test DVH (represented as a vector of Dx% values), the corresponding cluster was identified based on its Euclidean distance from the cluster centroids; a test DVH was considered to belong to a training cluster with the closest centroid. The proximity of a test DVH to cluster centroids is an objective way of deciding about the associated cluster, compared with visually identifying the dose-volume region that it falls in, because some DVHs may cross region borderlines. For each cluster (equivalently, dose-volume region) and PDE category, the actual incidence of ORN in the test set was calculated and compared against the estimated incidence based on the training risk indices and 95% CI. The estimates, for each cluster-PDE bin, were obtained by multiplying the training risk index (and 95% CI limits) by the total number of test patients in that bin. The results are presented in Tables 3 and 4. Except for one cluster-PDE bin, the actual ORN incidence numbers fell within the 95% CI of the training estimates for both K = 5 and K = 6. The exception was the bin of cluster No. 3 and PDE = 1 for K = 5, where the actual number of ORN was 4 (out of 24 patients in that bin), while the expected number (based on the training risk index) was 2.1 with the 95% CI of (0.3, 3.9) patients.

Discussion

The results show significant correlations between the risk of ORN and radiation intensity as well as PDE status, which is aligned with the univariate analysis results and confirms the relevance of structural clusters to the risk of ORN development. We note that cluster No. 6 (black) was identified as a part of cluster No. 5 (red) with K = 5, which implies that these 2 clusters are most similar structurally. The results show that these 2 clusters also have very similar risk indices in both PDE categories (25.6% vs 26.3% for PDE = 0 and 29.2% vs 29.8% for PDE = 1). In addition, the risk indices reported under K = 5 are very close to those under K = 6 with absolute deviation of less than 1% across dose-volume regions; see Appendix D (Supplementary Materials) for a direct comparison. These show consistency of the analysis results and further reveal the merit of (whole) DVH shape for predicting ORN.

The main advantage of the presented analysis is evaluation of radiation plans based on the whole DVH. As mentioned earlier, due to the multicollinearity issue, DVHs are reduced to a single-value summary in classification-based NTCP models, which disregards the fact that clinically different DVH curves—with different toxicity outcomes—may share a feature value. This can be observed by the toxicity outcome for the patients whose D30% is in the range of 55 Gy to 60 Gy, in our data set. According to van Dijk et al[11] model, the risk of ORN for these patients (per PDE status) is relatively the same, as their mandibular DVHs have close D30% values. However, the actual treatment outcomes for these patients are different and arguably determined by the shape of DVHs. Figure 5 shows these patients' DVHs; the incidence of ORN among the blue DVHs is 13.6% (3 out of 22) and among the black DVHs is 23.5% (12 out of 51). Other than this important difference, the van Dijk et al[11] and our models predict similar risks for these patients (ie, 13.0% to 19.0% by van Dijk et al[11] compared with 12.3% to 25.6% (of the blue and black clusters) by our model for PDE = 0 and 22.4% to 31.2% by van Dijk et al[11] compared with 23.6% to 29.2% (of the blue and black clusters) by our model for PDE = 1.

A significant advantage of our proposed method to address multicollinearity is its power to maintain the capacity for clinical interpretation. Although existing methods such as principal component analysis (PCA)[32,33] can resolve the issue of multicollinearity through synthesized features (which are obtained by transforming the original features), they fall short in maintaining interpretability, which renders them less favorable for clinical decision making.

Incidence rate is the unique minimum-variance unbiased estimator for the parameter of a Bernoulli random variable representing unmeasurable (hidden) features. In the absence of information about the main separating features (ie, highly inseparable classes by measurable features), such an estimator offers the most that can be learned from the data. This is supported by the close performances of logistic regression and naive Bayes classifiers on our data (area under the curve of 0.77 for logistic regression vs 0.75 for naive Bayes; see Appendix E Supplementary Materials). Naive Bayes is a simple classifier whose predictions mainly rely on the observed incidence rates.[29] Logistic regression generally performs better than naive Bayes, as the latter assumes independence among model features. The close performance of logistic regression and naive Bayes on our data indicates that, due to

inseparability of classes, the logistic regression predictions are mainly driven by ORN incidence rates with respect to the model features. Such estimates will be naturally more accurate if the data are divided into categories characterized by statistically significant (measurable) features. Thus, the presented analysis not only considers the whole DVH (not a single-value summary), but also has a simple yet robust statistical-inference foundation.

Although our model predictions are based on delivered dose (DVH) and PDE status, the univariate analysis of our data set has also identified tumor site ("cancer subsite" in the list of data features) as a statistically significant predictor of ORN. This feature, however, has been excluded from the final model by the Akaike Information Criterion-based forward stepwise selection, due to its high correlation with the dose features. This implies that tumor site is a similar yet weaker predictor of ORN and the delivered radiation dose captures its effect reasonably well. We also note that, depending on the tumor location, there might be little flexibility for organ sparing.

The proposed approach has its own limitations to be considered. First, the choice of the number of clusters (K) in the K-means method is relatively subjective. Although the inertia criterion provides straightforward guidance to identify the number of clusters from the mathematical perspective, our analysis showed that such a recommendation is not necessarily optimal from the application standpoint; repeating the analysis with a few different values of K and comparing and contrasting the results will lead to well-informed conclusions. Second, confidence on the risk indices rendered by our model directly relies on the number of observations for each cluster, independent of the others. For example, the inference of $r = 0.0\%$ for the first and second regions of the dose-volume space in Figure 4b is made based on 9 and 8 observations, respectively. Although this estimate is acceptable for these low-dose regions, in general, such small samples are highly sensitive to noise. ORN has low prevalence, which makes the classes of ORN and non-ORN patients unbalanced. This poses a challenge to statistical inference, in general, both supervised- and unsupervised-learning methods. Finally, the loss of spatial information about dose distribution is a shortcoming of all DVH-based predictive models; using the original 3D dose grid can lead to more accurate predictions.

The presented results, particularly the well-defined partitioning of the dose-volume space by DVHs, are attributed to the specific shapes of the mandible DVHs for ORN risk assessment. For other organs and toxicities, the DVHs may not behave in a similar fashion. Our data set exhibited a single overlapping cluster; however, this might not be the case in general, even for ORN data sets. Hence, the applicability of the presented method for risk assessment of other radiation toxicities needs to be investigated separately. We also encourage researchers to examine the performance of this method on other ORN data sets through external validation.

Visualization tools play a key role in clinical interpretability and reliability of mathematical models. In our analysis, visualization of the identified clusters by the K-means clustering method uncovered the distinction between the sixth and first 5 clusters and led to an improved partitioning of the dose-volume space. Visualization of the intermediate results may also provide additional insight for radiation planning. For example, the sixth (black)

cluster may be named the cluster of postoperative radiation plans; more than 77% of the patients in this cluster have received postoperative radiation therapy while this percentage for the other clusters remains below 20%. The special shape of this cluster is easily detectable to oncologists, which immediately reveals the relatively high prevalence of ORN among this group of patients. Availability of a visualization dashboard to illustrate every step of the process (Figs. 1–3)—rather than just the final result (Fig. 4)—can greatly improve clinical reliability of the proposed analysis for risk evaluation of ORN (with different data sets) and other complications.

Conclusion

This article presents an unsupervised-learning analysis method for mandibular ORN risk assessment and its results on retrospective data of a large cohort of patients with HNC treated at The University of Texas MD Anderson Cancer Center between 2005 and 2015. As the proposed method characterizes the ORN risk based on the whole DVH (in contrast to single-value summary as sought in existing data-driven NTCP models), the results provide a range of dose constraints (limit on various Dx% values for dose optimization) to ensure that the risk of ORN stays below a certain level.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Disclosures:

A.J.S. and C.D.F. received/receive related funding and salary support from National Science Foundation/National Institutes of Health National Cancer Institute (NCI) via the National Science Foundation/National Institutes of Health Smart and Connected Health Program (R01CA257814). C.D.F., A.S.R.M., and S.Y.L. received related support from the National Institute of Dental and Craniofacial Research awards U01DE032168, R56/R01DE025248, and R01DE028290. C.D.F. also received support from NCI under the MD Anderson Cancer Center Support Grant (P30CA016672) Image-Driven Biologically Informed Therapy Program. C.D.F. has received unrelated support through National Institutes of Health, National Science Foundation, and Patient-Centered Outcomes Research Institute Subaward. C.D.F. has received unrelated direct industry grant/in-kind support, honoraria, and travel funding from Elekta AB, and travel funding from Varian/Siemens Healthineers and Philips Medical Systems. C.D.F. has served in an unrelated consulting capacity for Varian/Siemens Healthineers, Philips Medical Systems, and Oncospace, Inc. C.D.F. receives grant and infrastructure support from MD Anderson Cancer Center via the Charles and Daneen Stiefel Center for Head and Neck Cancer Oropharyngeal Cancer Research Program and the Program in Image guided Cancer Therapy. S.Y.L. has received support from NCI awards R01CA280980, R21CA259839, and T32CA261856. S.Y.L. has served in a consulting capacity for Cardinal Health. A.S.R.M. has received support from NCI award R01CA258827. T.C.S. has received support from The University of Texas Health Science Center at Houston Center for Clinical and Translational Sciences TL1 Program (TL1TR003169).

Data Sharing Statement:

Research data are available at https://doi.org/10.6084/m9.figshare.13568207.

References

1. Jacobson AS, Buchbinder D, Hu K, Urken ML. Paradigm shifts in the management of osteoradionecrosis of the mandible. Oral Oncology 2010;46:795–801. [PubMed: 20843728]

2. Lee IJ, Koom WS, Lee CG, et al. Risk factors and dose-effect relationship for mandibular osteoradionecrosis in oral and oropharyngeal cancer patients. Int J Radiat Oncol Biol Phys 2009;75:1084–1091. [PubMed: 19327914]
3. Murray CG, Herson J, Daly TE, Zimmerman S. Radiation necrosis of the mandible: A 10-year study. Part I. Factors influencing the onset of necrosis. Int J Radiat Oncol Biol Phys 1980;6:543–548. [PubMed: 7410128]
4. Teng MS, Futran ND. Osteoradionecrosis of the mandible. Curr Opin Otolaryngol Head Neck Surg 2005;13:217–221. [PubMed: 16012245]
5. Tsai CJ, Hofstede TM, Sturgis EM, et al. Osteoradionecrosis and radiation dose to the mandible in patients with oropharyngeal cancer. Int J Radiat Oncol Biol Phys 2013;85:415–420. [PubMed: 22795804]
6. Chronopoulos A, Zarra T, Ehrenfeld M, Otto S. Osteoradionecrosis of the jaws: Definition, epidemiology, staging and clinical and radiological findings. A concise review. Int Dent J 2018;68:22–30. [PubMed: 28649774]
7. Marx RE. A new concept in the treatment of osteoradionecrosis. J Oral Maxillofac Surg 1983;41:351–357. [PubMed: 6574217]
8. Pitak-Arnnop P, Sader R, Dhanuthai K, et al. Management of osteoradionecrosis of the jaws: An analysis of evidence. Eur J Surg Oncol 2008;34:1123–1134. [PubMed: 18455907]
9. Aarup-Kristensen S, Hansen CR, Forner L, et al. Osteoradionecrosis of the mandible after radiotherapy for head and neck cancer: Risk factors and dose-volume correlations. Acta Oncol 2019;58:1373–1377. [PubMed: 31364903]
10. Sijtsema ND, Verduijn GM, Nasserinejad K, et al. Development of a local dose-response relationship for osteoradionecrosis within the mandible. Radiother Oncol 2023 109736. [PubMed: 37315578]
11. van Dijk LV, Abusaif AA, Rigert J, et al. Normal tissue complication probability (NTCP) prediction model for osteoradionecrosis of the mandible in patients with head and neck cancer after radiation therapy: Large-scale observational cohort. Int J Radiat Oncol Biol Phys 2021;111:549–558. [PubMed: 33965514]
12. Jereczek-Fossa BA, Orecchia R. Radiotherapy-induced mandibular bone complications. Cancer Treat Rev 2002;28:65–74. [PubMed: 12027415]
13. Otto S, Pautke C, Van den Wyngaert T, Niepel D, Schiødt M. Medication-related osteonecrosis of the jaw: Prevention, diagnosis and management in patients with cancer and bone metastases. Cancer Treat Rev 2018;69:177–187. [PubMed: 30055439]
14. Yang D, Zhou F, Fu X, et al. Symptom distress and interference among cancer patients with osteoradionecrosis of jaw: A cross-sectional study. Int J Nurs Sci 2019;6:278–282. [PubMed: 31508447]
15. Li XA, Alber M, Deasy JO, et al. The use and QA of biologically related models for treatment planning: Short report of the TG-166 of the therapy physics committee of the AAPM. Med Phys 2012;39:1386–1409. [PubMed: 22380372]
16. Marks LB, Yorke ED, Jackson A, et al. Use of normal tissue complication probability models in the clinic. Int J Radiat Oncol Biol Phys 2010;76:S10–S19. [PubMed: 20171502]
17. Palma G, Monti S, Conson M, Pacelli R, Cella L. Normal tissue complication probability (NTCP) models for modern radiation therapy. Semin Oncol 2019;46:210–218. [PubMed: 31506196]
18. Kutcher GJ, Burman C. Calculation of complication probability factors for non-uniform normal tissue irradiation: The effective volume method. Int J Radiat Oncol Biol Phys 1989;16:1623–1630. [PubMed: 2722599]
19. Stieb S, Lee A, van Dijk LV, et al. NTCP modeling of late effects for head and neck cancer: A systematic review. Int J Particle Ther 2021;8:95–107.
20. Bakhshandeh M, Hashemi B, Mahdavi SRM, et al. Normal tissue complication probability modeling of radiation-induced hypothyroidism after head-and-neck radiation therapy. Int J Radiat Oncol Biol Phys 2013;85:514–521. [PubMed: 22583606]
21. Beetz I, Schilstra C, Burlage FR, et al. Development of NTCP models for head and neck cancer patients treated with three-dimensional conformal radiotherapy for xerostomia and sticky

saliva: The role of dosimetric and clinical factors. Radiother Oncol 2012;105:86–93. [PubMed: 21632133]

22. Beetz I, Schilstra C, van der Schaaf A, et al. NTCP models for patient-rated xerostomia and sticky saliva after treatment with intensity modulated radiotherapy for head and neck cancer: The role of dosimetric and clinical factors. Radiother Oncol 2012;105:101–106. [PubMed: 22516776]
23. Buettner F, Miah AB, Gulliford SL, et al. Novel approaches to improve the therapeutic index of head and neck radiotherapy: An analysis of data from the PARSPORT randomised phase III trial. Radiother Oncol 2012;103:82–87. [PubMed: 22444242]
24. Dean JA, Wong KH, 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. [PubMed: 29399642]
25. 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;120:21–27. [PubMed: 27240717]
26. Morimoto M, Bijl HP, van der Schaaf A, et al. Development of normal tissue complication probability model for trismus in head and neck cancer patients treated with radiotherapy: The role of dosimetric and clinical factors. Anticancer Res 2019;39:6787–6798. [PubMed: 31810944]
27. van den Bosch L, van der Schaaf A, van der Laan HP, et al. Comprehensive toxicity risk profiling in radiation therapy for head and neck cancer: A new concept for individually optimised treatment. Radiother Oncol 2021;157:147–154. [PubMed: 33545258]
28. Wopken K, Bijl HP, van der Schaaf A, et al. Development of a multivariable normal tissue complication probability (NTCP) model for tube feeding dependence after curative radiotherapy/chemo-radiotherapy in head and neck cancer. Radiother Oncol 2014;113:95–101. [PubMed: 25443500]
29. James G, Witten D, Hastie T, Tibshirani R. An Introduction to Statistical Learning112Springer; 2013.
30. van Dijk LV, Fuller CD, Mohamed AS. Dose-volume histogram (DVH) parameters of the mandible for normal tissue complication probability modelling. Available at: 10.6084/m9.figshare.13568207. Accessed February 20, 2023.
31. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine learning in Python. J Mach Learn Res 2011;12:2825–2830.
32. Dean JA, Wong KH, Gay H, et al. Functional data analysis applied to modeling of severe acute mucositis and dysphagia resulting from head and neck radiation therapy. Int J Radiat Oncol Biol Phys 2016;96:820–831. [PubMed: 27788955]
33. Hansen CR, Bertelsen A, Zukauskaite R, et al. Prediction of radiation-induced mucositis of H&N cancer patients based on a large patient cohort. Radiother Oncol 2020;147:15–21. [PubMed: 32224314]

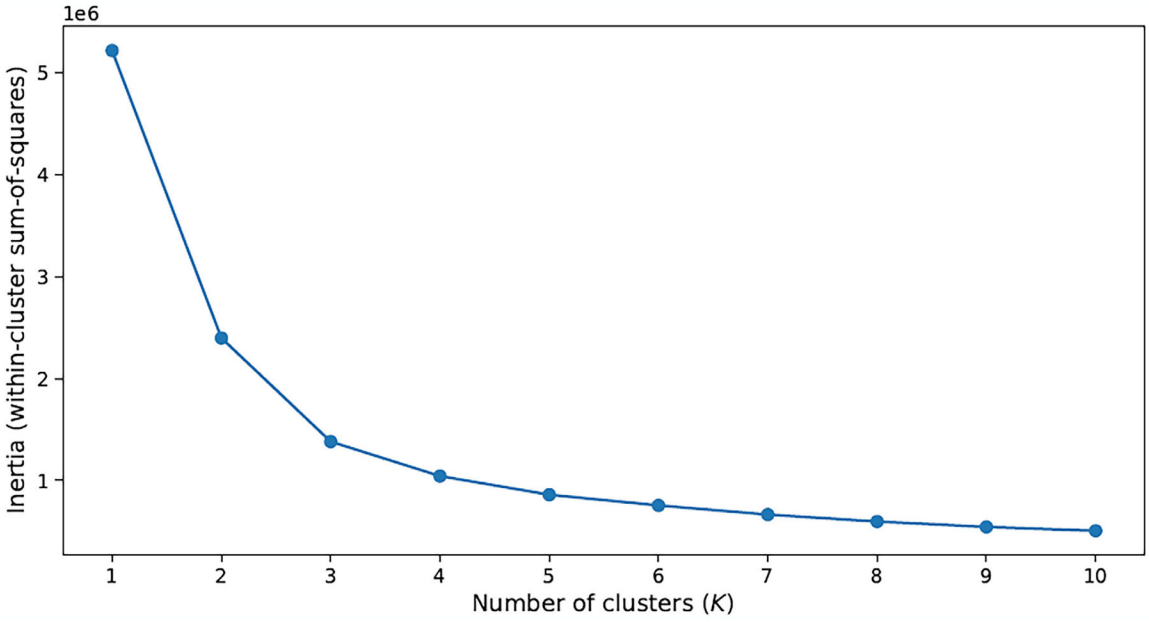

Fig. 1.
Inertia plot of the K-means clustering method.

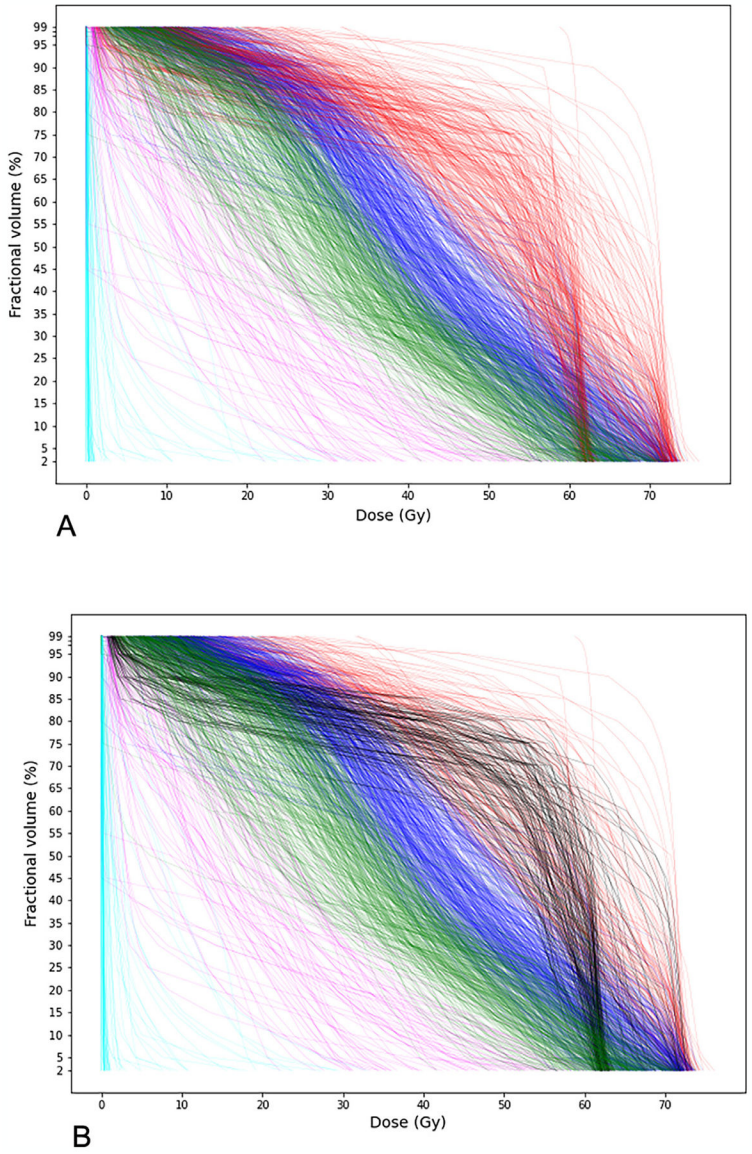

Fig. 2.
Visualization of the identified clusters. (A) K = 5; (B) K = 6.

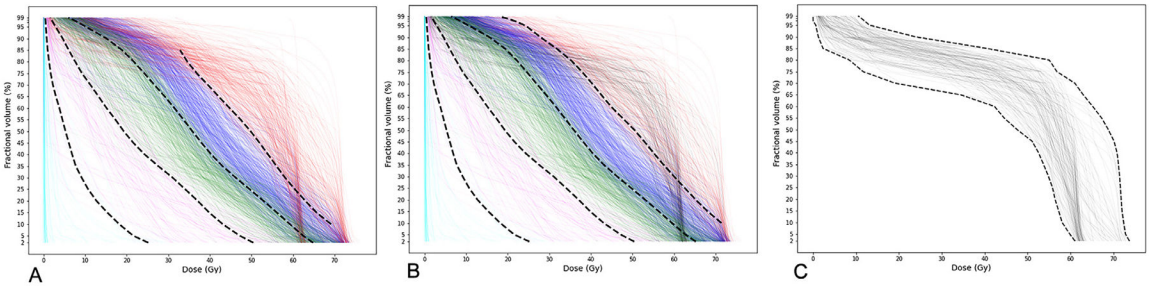

Fig. 3.

Cluster borders identified by soft-margin support vector machine. (A) K = 5: Cluster borders; (B) K = 6: Borders of the first 5 clusters; (C) K = 6: Envelopes of the sixth cluster.

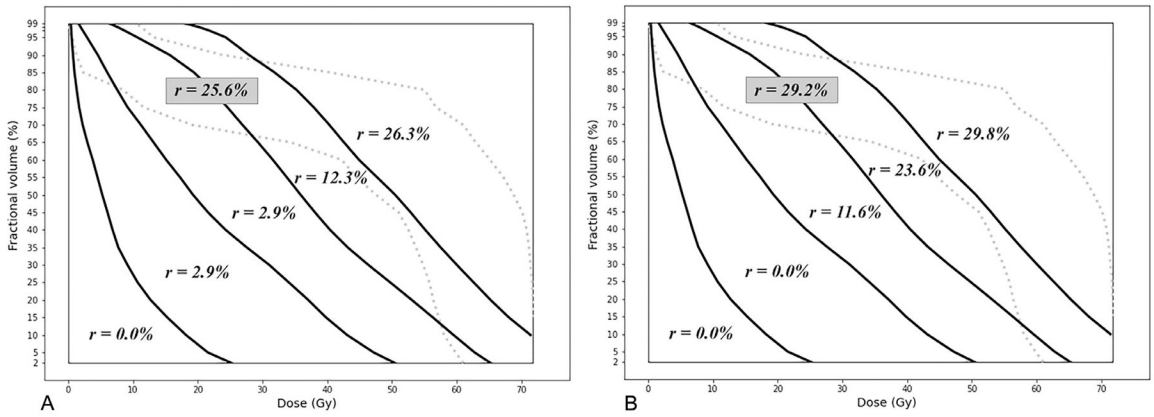

Fig. 4.

Risk indices of dose-volume regions for K = 6. (A) No/edentulous dental extractions (PDE = 0). (B) With dental extractions (PDE = 1).

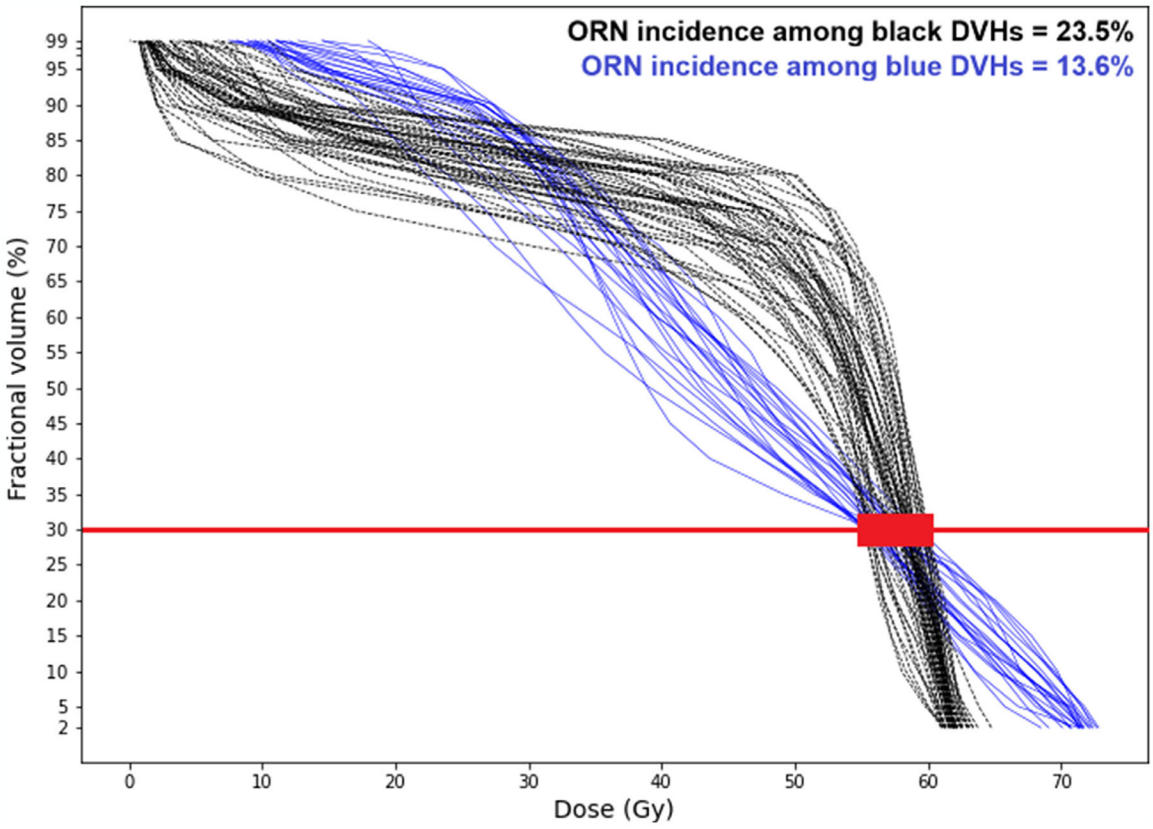

Fig. 5.

Different osteoradionecrosis incidences among dose-volume histograms with the same D30% value.

Table 1

ORN risk index (and 95% CI) for the identified clusters with K = 5 per PDE status

	PDE = 0		PDE = 1	
Cluster no.	**ORN incidence**	**Risk index (95% CI)**	**ORN incidence**	**Risk index (95% CI)**
1	0 out of 58	0.0%	0 out of 9	0.0%
2	2 out of 68	2.9% (0.0%, 6.9%)	0 out of 8	0.0%
3	8 out of 271	3.0% (1.0%, 5.0%)	9 out of 83	10.8% (4.1%, 17.5%)
4	41 out of 321	12.8% (9.1%, 16.5%)	37 out of 152	24.3% (17.5%, 31.1%)
5	50 out of 199	25.1% (19.1%, 31.1%)	26 out of 90	28.9% (19.5%, 38.3%)

Abbreviations: ORN = osteoradionecrosis; PDE = preradiation dental extraction (0 = no/edentulous, 1 = dental extractions).

Table 2

ORN risk index (and 95% CI) for the identified clusters with K = 6 per PDE status

	PDE = 0		PDE = 1	
Cluster no.	**ORN incidence**	**Risk index (95% CI)**	**ORN incidence**	**Risk index (95% CI)**
1	0 out of 58	0.0%	0 out of 9	0.0%
2	2 out of 68	2.9% (0.0%, 6.9%)	0 out of 8	0.0%
3	8 out of 273	2.9% (0.9%, 4.9%)	10 out of 86	11.6% (4.8%, 18.4%)
4	39 out of 318	12.3% (8.7%, 15.9%)	34 out of 144	23.6% (16.7%, 30.5%)
5	31 out of 118	26.3% (18.3%, 34.3%)	14 out of 47	29.8% (16.7%, 42.9%)
6	21 out of 82	25.6% (16.1%, 35.1%)	14 out of 48	29.2% (16.3%, 42.1%)

Abbreviations: ORN = osteoradionecrosis; PDE = preradiation dental extraction (0 = no/edentulous, 1 = dental extractions).

Table 3

Model validation: Estimated versus actual no. of ORN in the test set for K = 5 per PDE status

	PDE = 0				PDE = 1			
Cluster no.	**Training risk indices (95% CI)**	**Total no. of patients**	**Estimated no. of ORN (95% CI)**	**Actual no. of ORN**	**Training risk indices (95% CI)**	**Total no. of patients**	**Estimated no. of ORN (95% CI)**	**Actual no. of ORN**
1	0.0%	18	0.0	0	0.0%	2	0.0	0
2	4.4% (0.0%, 10.4%)	23	1.0 (0.0, 2.4)	0	0.0%	5	0.0	0
3	2.6% (0.3%, 4.9%)	73	1.9 (0.2, 3.6)	3	8.9% (1.4%, 16.4%)	24	2.1 (0.3, 3.9)	4
4	12.0% (7.8%, 16.1%)	92	11.0 (7.2, 14.8)	11	24.3% (16.3%, 32.3%)	43	10.5 (7.0, 14.0)	10
5	24.8% (17.5%, 32.3%)	70	17.4 (12.2, 22.6)	19	30.6% (19.2%, 42.2%)	29	8.9 (5.6, 12.2)	7

Abbreviations: ORN = osteoradionecrosis; PDE = preradiation dental extraction (0 = no/edentulous, 1 = dental extractions).

Table 4

Model validation: Estimated versus actual no. of ORN in the test set for K = 6 per PDE status

	PDE = 0				PDE = 1			
Cluster no.	**Training risk indices (95% CI)**	**Total no. of patients**	**Estimated no. of ORN (95% CI)**	**Actual no. of ORN**	**Training risk indices (95% CI)**	**Total no. of patients**	**Estimated no. of ORN (95% CI)**	**Actual no. of ORN**
1	0.0%	18	0.0	0	0.0%	2	0.0	0
2	4.4% (0.0%, 10.4%)	23	1.0 (0.0, 2.4)	0	0.0%	5	0.0	0
3	2.9% (0.4%, 5.4%)	66	1.9 (0.2, 3.6)	3	10.6% (1.8%, 19.4%)	18	1.9 (0.3, 3.5)	3
4	9.6% (5.8%, 13.4%)	90	8.6 (5.2, 12.0)	8	20.6% (12.8%, 28.4%)	45	9.3 (5.8, 12.8)	10
5	28.6% (17.5%,39.7%)	25	7.1 (4.3, 9.9)	8	36.4% (16.3%,56.5%)	9	3.3 (1.5, 5.1)	2
6	22.6% (14.1%, 31.1%)	54	12.2 (7.6, 16.8)	14	29.3% (17.6%, 41.0%)	24	7.0 (4.2, 9.8)	6

Abbreviations: ORN = osteoradionecrosis; PDE = preradiation dental extraction (0 = no/edentulous, 1 = dental extractions).