



## Precision association of lymphatic disease spread with radiation-associated toxicity in oropharyngeal squamous carcinomas



Andrew Wentzel<sup>a,\*</sup>, Timothy Luciani<sup>a</sup>, Lisanne V. van Dijk<sup>b</sup>, Nicolette Taku<sup>b</sup>, Baher Elgohari<sup>b,e</sup>, Abdallah S.R. Mohamed<sup>b</sup>, Guadalupe Canahuate<sup>d</sup>, Clifton D. Fuller<sup>b</sup>, David M. Vock<sup>c</sup>, G. Elisabeta Marai<sup>a,\*</sup>, Spatial-Non-spatial Multi-Dimensional Analysis of Radiotherapy Treatment/Toxicity Team SMART3

<sup>a</sup> Department of Computer Science, The University of Illinois at Chicago, Chicago; <sup>b</sup> Department of Radiation Oncology, The University of Texas MD Anderson Cancer Center, Houston; <sup>c</sup> Department of Clinical Oncology and Nuclear Medicine, Mansoura University, Mansoura, Egypt; <sup>d</sup> Department of Electrical and Computer Engineering, University of Iowa, Iowa City, USA; <sup>e</sup> Division of Biostatistics, University of Minnesota, Minneapolis

### ARTICLE INFO

#### Article history:

Received 16 October 2020  
Received in revised form 18 May 2021  
Accepted 5 June 2021  
Available online 11 June 2021

#### Keywords:

Oropharynx cancer  
Precision medicine  
Statistical data mining  
Medical informatics  
Radiation-associated dysphagia  
Head and neck cancer

### ABSTRACT

**Purpose:** To determine whether patient similarity in terms of head and neck cancer spread through lymph nodes correlates significantly with radiation-associated toxicity.

**Materials and methods:** 582 head and neck cancer patients received radiotherapy for oropharyngeal cancer (OPC) and had non-metastatic affected lymph nodes in the head and neck. Affected lymph nodes were segmented from pretreatment contrast-enhanced tomography scans and categorized according to consensus guidelines. Similar patients were clustered into 4 groups according to a graph-based representation of disease spread through affected lymph nodes. Correlation between dysphagia-associated symptoms and patient groups was calculated.

**Results:** Out of 582 patients, 26% (152) experienced toxicity during a follow up evaluation 6 months after completion of radiotherapy treatment. Patient groups identified by our approach were significantly correlated with dysphagia, feeding tube, and aspiration toxicity ( $p < .0005$ ).

**Discussion:** Our results suggest that structural geometry-aware characterization of affected lymph nodes can be used to better predict radiation-associated dysphagia at time of diagnosis, and better inform treatment guidelines.

**Conclusion:** Our work successfully stratified a patient cohort into similar groups using a structural geometry, graph-encoding of affected lymph nodes in oropharyngeal cancer patients, that were predictive of late radiation-associated dysphagia and toxicity.

© 2021 Elsevier B.V. All rights reserved. Radiotherapy and Oncology 161 (2021) 152–158

Ever larger numbers of head and neck cancer (HNC) patients survive years after oncologic therapy due to increased efficacy of therapy, increased incidences of human papilloma virus (HPV) related HNC, and decreased numbers of smoking and tobacco related tumors [1]. Though they may survive therapy, patients are often plagued with long lasting or permanent residual toxicity. We hypothesize that information about a patient's pattern of disease spread through lymph nodes could be used as a prognostic indicator for symptoms of late Radiation Associated Dysphagia (RAD) toxicity. To evaluate this hypothesis, we developed a novel patient risk stratification method based on grouping HNC patients with similar patterns of disease spread through the lymph nodes,

and applied it to imaging data collected retrospectively from a cohort of 582 HNC patients.

Concretely, we encoded chains of affected lymph nodes as a set of covariates that incorporate spatial relationships between affected nodes, used this representation to calculate patient similarity, and applied unsupervised clustering to stratify patients into 4 groups. We show that the resulting groups are strongly correlated with two post-treatment toxicities associated with RAD. These groups serve as a precise toxicity-risk staging system based on nodal disease spread.

Despite an overall reduction in tobacco and alcohol associated malignancies of the head and neck cancers (HNC), recent decades have been marked by a paradoxical increase in the incidence of cancers of the oropharynx (OPC) [2]. Approximately 7500 cases of OPC are diagnosed in the United States annually, including 70–90% of which are associated with the human papilloma virus (HPV) [3–5]. HPV-related OPC patients generally have better prognosis than HPV-negative OPC patients, with 5-year overall survival

\* Corresponding authors at: Department of Computer Science, The University of Illinois at Chicago, Chicago, IL 60612, USA.

E-mail addresses: [awentze2@uic.edu](mailto:awentze2@uic.edu) (A. Wentzel), [gmarai@uic.edu](mailto:gmarai@uic.edu) (G. Elisabeta Marai).

<sup>1</sup> Questions about statistical analyses can be directed to Andrew Wentzel.

rates around 80%, compared to rates of less than 50% for HPV-negative patients [6]. The steady increase of OPC survivors has amplified the need to understand and minimize acute and long term side effects following (chemo-)radiation based on cohorts of similar patients [7,8]. Swallowing dysfunction, i.e. radiation-associated dysphagia (RAD), has a particularly large impact on the quality of life of OPC survivors [9]. RAD can result in nutritional deficiencies and life-threatening aspiration pneumonia [10]. Over 50% of patients with locally-advanced OPC will demonstrate acute or subacute evidence of aspiration and nearly 10% will become gastrostomy tube dependent during the peri-treatment course [11–14].

Studies have shown that the maximum distance between involved lymph nodes and primary tumors are potential determinants of metastasis-free survival, and that patterns of involved (diseased) lymph nodes (LN), as well as tumor proximity to organs at risk, may affect toxicity [15,16]. However, matching spread patterns based on lymph labels alone and ignoring spread over adjacent anatomical regions has been shown to result in incorrect patient matches [17]. These data suggest that spatial information pertaining to lymph node involvement may be useful in predicting patient outcomes.

Given the overall favorable survival outcomes of HPV-associated disease, there is a need for staging tools that can integrate baseline clinical information and stratify patients by risk for treatment-related clinical toxicities. By identifying patients at higher risk of complications, physicians can identify where to recommend treatment changes or preventative therapy to maintain swallowing function during and after treatment. For example, moderate risk patients could be given more aggressive swallowing exercises, nutritional support, or post-therapy surveillance for dysphagia symptoms. High risk patients could also be selected for proton therapy referral or given prophylactic gastrostomy tube placement. As OPCs typically exhibit substantial lymphatic involvement, pre-therapy predictive models for toxicity that are available before radiation planning are exceptionally valuable, as they can allow for actionable dose-modification or surgical neck management.

**Material and methods**

*Patient cohort*

An IRB-approved, retrospective review of patients diagnosed with oropharyngeal cancer (OPC) and treated at MD Anderson Cancer Center from 2005 to 2013 was performed. Patients with lymph node positive, biopsy-proven, OPC who were treated with radiotherapy (RT) with or without chemotherapy with curative intent and had at least 6 months post-treatment follow-up assessment were eligible for inclusion. Patients’ demographic, clinical, treatment, and outcome measures are summarized in (Table 1).

Patient demographics, clinical data, outcomes, and toxicity assessments were collected retrospectively from electronic medical records. All patients underwent a complete physical and endoscopic examination, as well as radiological and pathological assessments at initial treatments. Data on affected lymph nodes levels were taken from baseline contrast-enhanced tomography (CECT) scans obtained during initial assessment. Lymph node levels extracted from CECT scans were used to construct a map of each patient’s lymphatic disease spread. LN levels were defined based on anatomical landmarks. LN levels considered were the retropharyngeal larynx (RPN) and levels I-VI based on consensus guidelines [18]. Outcome data was collected during a 6-month follow-up assessment that included physical, endoscopic, radiological, nutritional and toxicity assessments. Radiation associated dysphagia was defined as the presence of grade 2+ aspiration rate

**Table 1**

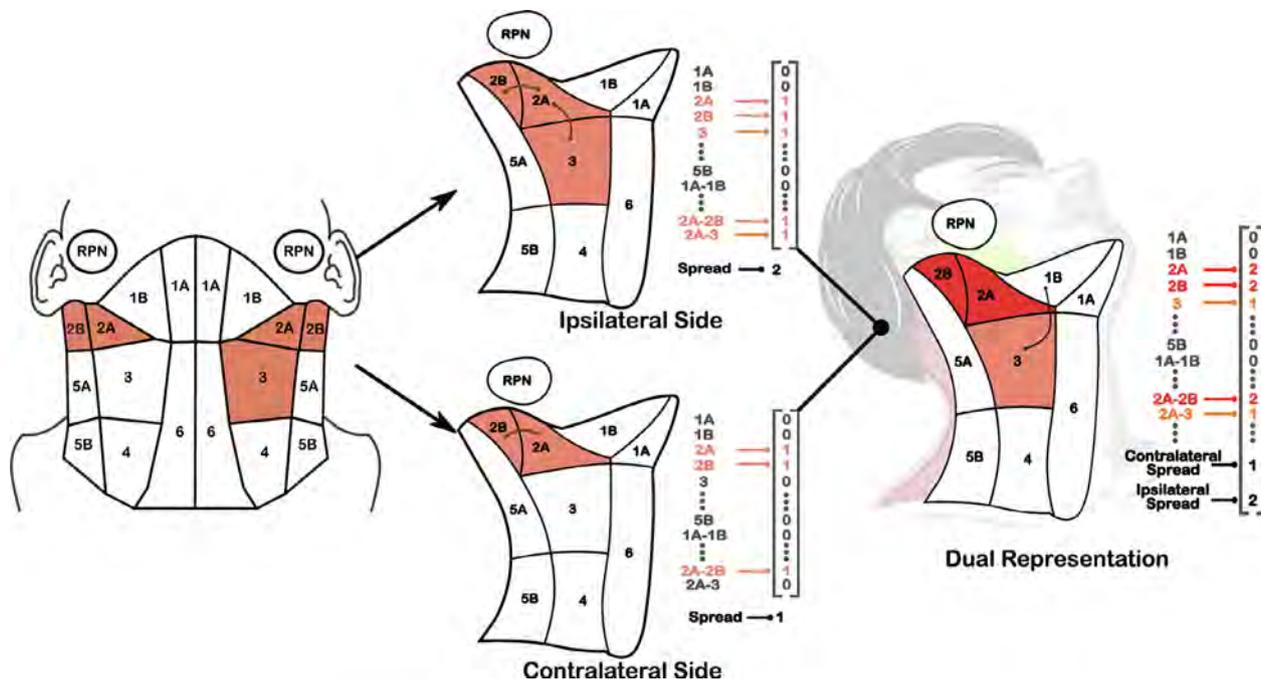
Patient Characteristics. Legend: IQ: Inter-quantile, NOS: not otherwise specified, GPS: glossopharyngeal sulcus, RT: radiation therapy, IC: induction chemotherapy, CC: concurrent chemo-radiation, Gy: gray. Intra-cluster breakdowns of pathological characteristics are available in Appendix B.

Characteristic	N (%)
<b>Gender</b>	
Male	512 (88)
Female	70 (12)
<b>Median age at diagnosis (IQ Range)</b>	<b>57.8 (52.1–65)</b>
<b>Race</b>	
White/Caucasian	530 (91.1)
Black/African American	17 (2.92)
Hispanic/Latino	20 (2.92)
Asian	7 (1.2)
Native American	1 (0.66)
NOS	7 (1.2)
<b>Smoking status</b>	
Current	118 (20)
Former	216 (37)
Never	248 (43)
<b>HPV Status</b>	
Negative	45 (8)
Positive	360 (62)
Unknown	177 (30)
<b>Tumor Sub-site</b>	
Tonsil	237 (41)
BOT	295 (51)
Soft palate	6 (1)
GPS	11 (2)
NOS	33 (6)
<b>Therapeutic combination</b>	
Radiation alone (RT alone)	74 (13)
Concurrent chemo-radiation (CC)	308 (53)
Induction chemotherapy (IC) + RT alone	53 (9)
IC + CC	147 (25)
<b>Total RT dose, median (IQ Range)</b>	<b>70 (66–70) Gy</b>
<b>Total RT fractions, median (IQ Range)</b>	<b>33 (30–33)</b>
<b>Feeding tube at 6 months</b>	
Y	99 (17)
N	483 (83)
<b>Aspiration rate at 6 months</b>	
Y	96 (16)
N	486 (84)
<b>N Category (7th Edition)</b>	
N1	71 (12)
N2a	49 (8)
N2b	311 (53)
N2c	133 (23)
N3	18 (3)
<b>T-Category</b>	
T1	131 (23)
T2	245 (42)
T3	121 (21)
T4	85 (15)

based on CTCAE guidelines [19] or feeding-tube insertion during treatment or after treatment completion. No feeding tubes were placed prophylactically.

*Patient stratification*

To perform stratification, each patient’s chain of affected lymph nodes was encoded as a multidimensional vector (Appendix A). First, a map of levels I-VI, and the RPN, were constructed such that anatomically adjacent regions were connected in the map (Fig. 1). Involved lymph nodes in each patient were identified for both sides of the head through physical examination and radiological imaging [20,21]. If at least one lymph node in a given level is affected with cancer cells, we consider the corresponding level to be “involved”. Fig. 1 illustrates the encoding process for an example patient with bilaterally affected levels IIA-IIB and unilaterally affected level 3.



**Fig. 1.** (Left) Diagram of the anatomically-adjacent representation of affected lymph nodes, for an example patient with levels 2A-2B affected on both sides of the head and level 3 affected on the ipsilateral side of the head. Adjacent regions are represented as bigrams (e.g., 2A-2B) in our encoding (Appendix A). (Center) Encodings for each side of the head. The contralateral side is given a 'maximum spread' value of 2 to represent the travel distance from 2b to 3, while the contralateral side (opposite primary tumor) has a spread of 1 for 2A-2B. (Right) Composite encoding for the patient that captures symmetry of the head and neck; diseased regions are added while the maximum LN spread on each side are treated as two separate covariates.

Pairwise similarity between patients was calculated using similarity based on LN spread over adjacent anatomical regions, which is an extension of the methodology introduced by Luciani et al. [17]. Hierarchical agglomerative clustering (HAC) was performed on the set of unique patterns of involvement present in the cohort. Patients were then assigned to clusters by matching their lymph node involvement to the cluster with the corresponding pattern of involvement. More details on the similarity calculation and clustering method is given in (Appendix A). Predictive models of toxicity 6-months after finishing radiation treatment were created using different numbers of clusters (2–5), with and without inclusion of N-staging. Baseline models were also created using only N-staging in order to compare the performance of the clusters to N-staging alone. For these models, membership in each cluster was encoded as a set of 4 binary variables. The optimal number of clusters was chosen using 10-fold cross validation and a logistic regression model for toxicity, described below. Statistical analysis and a descriptive characterization was performed on the groups. Reported clusters are labeled in the order of percentage of patients with RAD, where cluster 4 had the highest risk.

### Statistical analysis

Correlations between the LN cluster labels and toxicity outcome (i.e. RAD) and other clinical covariates (e.g. TNM staging, HPV status, GTVp dose, tumor subsite, age groups) were evaluated using statistical significance tests (Fisher's and Pearson Chi-square test) [22]. The likelihood ratio test was used to evaluate the performance of a logistic regression model for RAD using clinical covariates with and without including the LN cluster labels as covariates. A more detailed description of the methods used are given in (Appendix B).

Logistic regression models were created in order to compare the performance of different numbers of clusters with and without including N-staging, as well as to compare to the performance of

a baseline model where all patients are in a single cluster. These models were evaluated using a 10-fold cross-validation scheme where the model was trained on 90% of the data and the AUC was evaluated using the predicted risk for the remaining 10% of the data. This was repeated 10 times using separate splits, and the mean AUC score across all 10 splits is reported as the overall AUC score.

The likelihood ratio test (LRT) was used to show that the LN cluster labels were significantly correlated with RAD while accounting for other clinical factors, including age, HPV status, N-stage, T-stage, and treatment modality. Tumor subsite was not considered as it was not significantly correlated with RAD ( $p > .01$ ). The likelihood ratio test (LRT) is not a performance metric like the C-index/ROC or R2 which is prone to overoptimism due to overfitting. The likelihood ratio is a test of whether additional model terms are significantly different from zero and accounts for the fact that additional terms lead to an increased log likelihood.

For all regression models, cluster membership and clinical staging was encoding as categorical binary variables, where the cluster or clinical stage each patient was assigned to was encoded as a 1 and all other clusters were encoded as a 0. A more detailed description of the methods used for statistical methods are given in (Appendix B).

LN graph calculations and patient similarity computations were performed in Python using the pandas library [23]. Hierarchical clustering and statistical tests were performed using Matlab 2018a statistical toolbox and R. All statistical tests were two-sided with  $p < 0.01$  considered statistically significant.

### Results

Out of 644 OPC patients available for the study, 582 patients had affected lymph nodes and were included in the final cohort. Ignoring laterality, 63 distinct patterns of affected LNs were pre-

sent in the 582 patient cohort, with 6 patterns comprising 78% of the cases (Table 2). Of the patterns with over 10 cases, patients with bilaterally affected 2A-2B-3 levels showed the highest incidence of RAD (53%), while the group with unilaterally affected 2A-2B-3 regions had the lowest incidence (18%). Levels 2A-2B were the most common affected sites, with 76% and 19% of patients showing unilateral spread and bilateral involvement, respectively.

Table 3 shows the mean AUC for predicting feeding tube dependency, aspiration, and RAD using cluster labels for 2–5 clusters using 10-fold cross-validation. All models that included both N-staging and LN clusters performed better than the baseline model with only N-stage. In the subsequent analysis, we used 4 clusters, as they performed best for overall dysphagia (RAD) in terms of AUC, and facilitated direct comparison against the current AJCC T and N staging systems, which use 4 categories each to describe the size of the tumor and spread to lymph nodes.

The 4 groups consist of 1 large, low risk cluster (Cluster 1), and 3 smaller, higher risk clusters that are characterized by the spread and presence of bilaterally affected regions. To better explain the patient stratification to clinicians, we created a visual representation of the disease spread patterns in the clusters (Fig. 2) [24] using visual communication rules [25].

The 4 clusters were significantly correlated ( $p < .0001$ ) with both feeding tube and aspiration rate at 6 months post-therapy (Appendix C). Although clusters were also correlated ( $p < .0001$ ) with several other early risk factors: N-category, AJCC staging, T-category, tumor laterality, and total dose, none of these factors by themselves were able to stratify the patient cohort as our approach did. Clusters were also significantly correlated with HPV status ( $p < .01$ ). No correlation was found between clusters and age ( $p > .05$ ), nor with tumor subsite ( $p > .05$ ). A detailed discussion of each cluster is given in (Appendix C) and breakdowns of demographics and toxicity by cluster are detailed in (Appendix C Table C1).

After adjusting for age, smoking status, use of chemotherapy, HPV status, T category, and N category, LN clusters were significantly associated with RAD using a likelihood ratio test ( $p < 0.001$ ). The addition of the LN clusters shows significant improvement when included in the model, even when T and N-category are included (Table 4).

## Discussion

Head and neck cancers account for nearly 3% of all malignancies in the U.S. with approximately 62,000 HNC cases diagnosed per year [3]. More than two-thirds of those diagnosed with HNC will survive 5 years or more if treated with locoregional curative therapy. However, almost all radiotherapy survivors will suffer from at least mild-to-moderate symptoms from head and neck radiation [26,16]. Recent phase III studies [27,28] suggest that concurrent chemoradiation will remain the standard for most locoregional head and neck cancers, and treatment protocols that minimize toxicity risk will remain essential. Thus, effective methods to predict radiation effects before radiation planning, actively reducing the side effects to adjacent organs-at-risk (OARs) deliverable with current radiotherapy treatment devices represents a clinically key improvement in cancer care. A majority of models utilized for toxicity prediction consider organs-at-risk (OARs) independently, but fail to consider distinct patterns of lymphatic spread, or that OAR dose is a function of proximity to tumor volumes [29].

We introduce a stratification technique to group patients based on the spread of diseased lymph nodes in OPC patients. These groups strongly correlated with late Radiation-Associated Dysphagia (RAD) after correcting for existing diagnostic information. Since our model relies on stratification using unsupervised clustering, it

**Table 2**

Affected lymph-node (LN) patterns and toxicity outcomes. Percentages of FT and AS are given as a percentage of the total patients with a given pattern. Legend: FT: Feeding tube. AS: Aspiration. 2A, 2B, 3, 4: lymph node level 2a, ab, 3 and 4.

Pattern	Patient Count	FT	AS
Unilateral 2A-2B	227	24 (11%)	26 (11%)
Unilateral 2A-2B-3	125	16 (13%)	13 (10%)
Bilateral 2A-2B	37	10 (27%)	4 (11%)
Unilateral 2A-2B-3-4	28	9 (32%)	7 (25%)
Bilateral 2A-2B, Unilateral 3	19	6 (32%)	4 (21%)
Bilateral 2A-2B-3	17	5 (29%)	7 (41%)
Top 6 patterns total	453	70 (15%)	61 (13%)
Other (57) patterns	129	29 (22%)	25 (27%)
All 63 patterns	582	99 (17%)	96 (16%)

can allow for granular risk predictions for small populations of patients with uncommon or unique lymph node involvement. Thus, even before time intensive radiation plan optimization, an indication of the expected RAD can be estimated. Conclusively, this work shows that lymph node spread, determined at the time of diagnoses, can supply a clinically usable RAD risk assessment and guide toxicity-informed treatment decision making.

Our model does not rely on the delineation of organs-at-risk or radiation dose optimization, which are time expensive, but instead uses the lymph-node spread as a precursor for late toxicity. While our clusters are largely correlated with N-staging, we identified sub-patterns within these groups that are not captured by current staging considerations. Most of the difference in the cohort presented as either in N-category 2b or lower (low risk) vs N-category 2c or higher (high-risk). However, our 3 high risk clusters were able to identify groups of higher risk patients within and across N-category 2b and 2c that allowed for more granular risk prediction not captured by N-category alone. Furthermore, after adjusting for age, smoking status, use of chemotherapy, HPV status, T-category, and N-category, the LN clusters were still significantly associated with RAD. Our technique uniquely considers 3-d anatomy of the lymphatic system in the oropharynx, allowing us to capture more granular spatial information beyond what is captured by current staging systems.

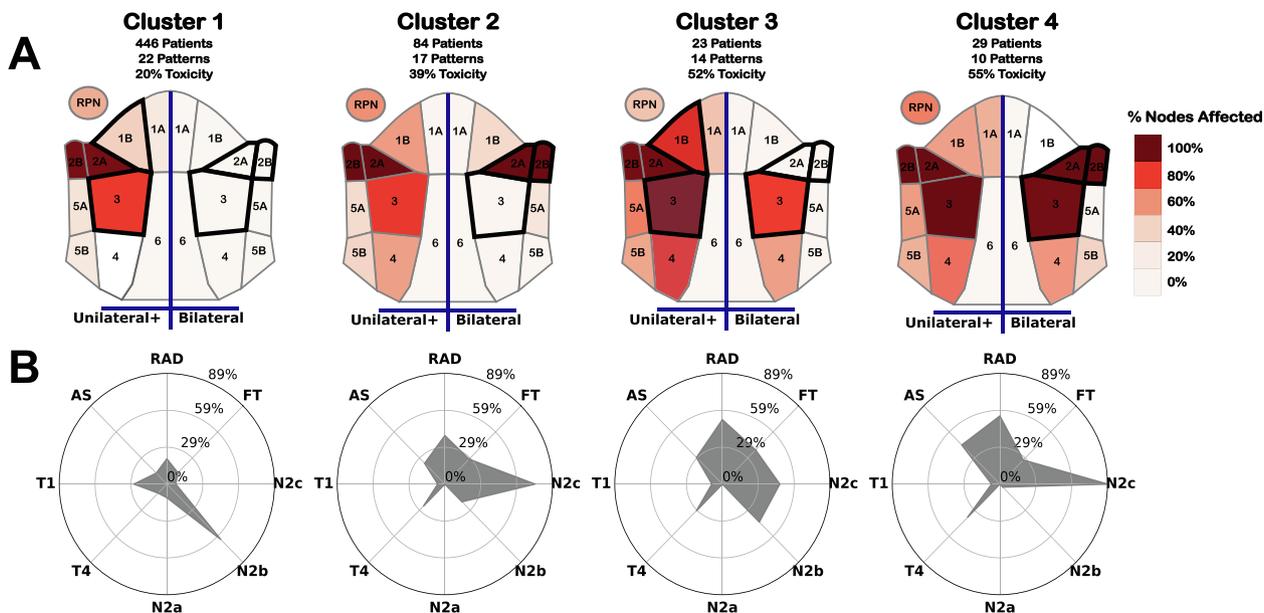
Interestingly, we found a disparity in risk within different clusters when comparing aspiration rate and feeding tube. Cluster 4 showed the highest percentage of RAD and a very high percentage of patients with aspiration. This is consistent with the group's higher degree of bilaterally affected nodes and larger nodal spread than other nodes, as well as the higher T-category and AJCC stage, as patients with more severe disease are more likely to have adverse effects. However, feeding tube toxicity was most prevalent in cluster 3, suggesting that some underlying causes of feeding tube are not captured by overall disease spread or traditional AJCC staging. When ignoring bilaterality, cluster 3 had the highest percentage of patterns with affected nodes in levels 1B, 4, 5A and 5B, indicating that cluster 3 may be separating a small set of patients with disease more localized to the lower regions of the neck, rather than the majority of patients with disease around levels 2A-2B. It may be the case then, that the presence of diseased lymph nodes in regions 4-5A-5B is indicative of feeding-tube toxicity risk separate from general disease spread or tumor size.

Our results suggest that bilaterality, as well as spread to level 3 or 4, are correlated with clusters at high-risk of RAD, and are thus an indicator of patient risk of toxicity. However, the fact that our clusters, while heavily correlated with RAD, are largely subsumed in N-category 2b/2c, suggests that existing classifications methods are insufficient for capturing relevant spatial information. We show that the maximum lymph node spread is a valuable covariate for stratifying patients, which is consistent with previous findings that suggest lymph node spread is a good measure of distance metastasis free survival [15]. Finally, some distinct patterns are

**Table 3**

AUC scores (10-fold cross validation) for toxicity prediction using logistic regression from different numbers of lymph node clusters, with and without addition of N-staging. Models with the highest AUC for a given outcome are in bold. AUC performance is highest for 3 clusters and for 4 clusters, implying that those provide the best ranking of patients based on risk prediction.

# Clusters	Cross-Validation AUC					
	Feeding Tube		Aspiration		RAD	
	Our method	+ N Category	Our method	+ N Category	Our method	+ N Category
Baseline (no clusters)	-	0.555	-	0.659	-	0.64
2 clusters	0.577	0.572	0.591	0.675	0.591	0.646
3 clusters	<b>0.612</b>	<b>0.592</b>	0.603	0.675	0.603	0.655
4 clusters	0.595	0.567	<b>0.621</b>	<b>0.677</b>	<b>0.621</b>	<b>0.656</b>
5 clusters	0.597	0.597	0.62	0.669	0.62	0.652



**Fig. 2.** Visual representation of each cluster. (A) Heat map of nodal spread within each cluster. Left half of each map indicates % of patients in the cluster with at least one affected node on a specific level, while the right half encodes the percentage of patients with bilateral spread within a specific level. Regions outlined in black denote regions that are most discriminative for cluster membership, and could be used to determine if 99% of patients are within a given cluster. (B) Radar chart showing the % of patients in each cluster with a given toxicity or inclusion in a specific clinical staging category. FT: Feeding Tube, AS: Aspiration, RAD: Radiation-associated dysphagia, T1-2: T-category 1 or 2, T3-T4: T-category 3 or 4, N0-N1: N-category 0 or 1, N2a-N2b: N-category 2a or 2b, N2c-N3: N-category 2c or 3.

**Table 4**

Akaike information criteria (AIC) and p-values from Likelihood ratio tests (LRT) for cluster labels in logistic regression models for predicting feeding tube, aspiration, and RAD. Lower AIC indicates better fit models after accounting for the number of covariates. LRT p-values <.05 indicates 95% confidence that cluster labels are meaningfully correlated with toxicity after accounting for other covariates. Legend: Age: Age >= 65; Smoke: Smoking Status (Never or Former/Current); Chemo: (Concurrent Chemo-radiation or Radiation Alone); HPV: HPV status (Positive, Negative, or Unknown); T: T-category (1-2 or 3-4); N: N-category (0-1 or 2-3).

Model Covariates	Feeding Tube			Aspiration Rate			RAD		
	Base model	Base model + LN Cluster		Base model	Base model + LN Cluster		Base model	Base model + LN Cluster	
	AIC	AIC	LRT	AIC	AIC	LRT	AIC	AIC	LRT
Clinical (Age, Smoke, Chemo, HPV)	519.1	508.6	<0.001	503.6	493.1	<0.001	645.1	627.1	<0.001
Clinical + T-stage	488.5	485	<0.05	477.2	472.3	<0.05	602.8	594.4	<0.01
Clinical + T-stage + N-stage	490.5	<b>486.8</b>	<0.05	475.2	<b>467.5</b>	<0.01	602.2	<b>590.9</b>	<0.001

more correlated with aspiration rate vs feeding tube toxicity, and while bilaterality, lymph-node spread, and T-category are good predictors of aspiration risk, the location of affected lymph node chains lower in the neck may be underexplored indicators of late feeding tube insertion, which is less well understood.

Since clustering was performed over the unique patterns of involvement, not the individual patients in the training cohort, assigning new patients to a cluster can be done without retraining the model. In the case that the new patient exhibits one of the patterns within the training data, the cluster assignment is done by

matching. In the case the new patient has a previously unseen pattern of involvement, the distance between the pattern and each cluster is computed and the patient is assigned to the closest cluster using average Euclidean distance.

While this retrospective study considers a cohort with many HNC patients and lymph node spread, it is not without limitations. First, our cohort consists primarily of white, male patients, and thus these results may not be reflected in more diverse demographics. The AJCC and TNM staging used in our analysis is the 7th edition, which was the standard at the time of treatment,

because of the large number of patients with unknown HPV status. However, HPV status, alongside TNM staging and treatment modality, is already considered in our multivariate logistic regression models when assessing statistical significance. In AJCC's 8th edition the N2a and N2b stage for HPV positive patients are combined with N1, as it provides a larger pool for events by grouping disparate lymph node distributions for overall survival risk cohort stratification. Comparatively, this work is concerned with late toxicity and its relation to LN involvement. The loss of granularity in the 8th edition would limit our ability to distinguish between lymph node patterns in HPV positive N2a and N2b patients.

Finally, we consider only patterns present in our cohort. While this assumption is reasonable for a majority of patients, there may be different factors that may indicate RAD risk in patients with uncommon patterns of nodal spread that are not present in our cohort (e.g., patients with nodal spread in level 6). Our results suggest that there is a small set of rare patterns with limited bilateral spread with higher risk of feeding tube toxicity. It is thus possible that there are spread patterns with bilaterally affected nodes in levels 3–5 that may be very high indicators of feeding tube insertion, but these patterns are rare, making it difficult to draw conclusions.

This work represents the first demonstration of a toxicity model incorporating spatial, inter-patient similarity metrics to predict radiation-associated dysphagia at 6-months post-radiation in oropharyngeal cancer survivors. Future efforts include work to apply these pre-trained clusters, alongside recent advances in delta-radiomics features, and radiomics-based clusters, to a different validation cohort, as well as adding normal tissue radiomics metrics to the spatial analyses presented herein. At present, while we have radiomics signatures associated with salivary injury, the data on muscle-specific radiomics markers of dysphagia remains preliminary but is a potential avenue for further hypothesis generation. Additionally, we seek to integrate spatial features into visualization software for dynamic decision tools [7] for multidisciplinary head and neck risk stratification, therapy selection, and patient education.

## Conclusion

In conclusion, this study demonstrates that clustering based on lymph nodes spread is associated with radiation-associated dysphagia, both aspiration toxicity and feeding tube dependency. Our anatomical representation of lymph node spread showed superior association to RAD compared to the current N-category classification. Our method relies only on discrete information on nodal spread at time of diagnosis, and thus does not require complex dose-planning or organ segmentation to determine RAD risk.

## Co-author specific contributions

All listed co-authors performed the following:

1. *Substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work;*
2. *Drafting the work or revising it critically for important intellectual content;*
3. *Final approval of the version to be published;*
4. *Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.*

Specific additional individual cooperative effort contributions to study/manuscript design/execution/interpretation, in addition to all criteria above are listed as follows:

- *AW, TL, GEM - designed and developed similarity measure, data extraction and curation, statistical analysis, and interpretation*
- *LVD, BE, ASRM, CDF - direct patient care provision, direct clinical data collection; interpretation and analytic support*
- *GC - supervised statistical analysis, data extraction, and assisted with interpretation*
- *DV, CDF - analytic support, guarantor of statistical quality*
- *AW, GC, LVD, ASRM, NT, CDF, GEM - manuscript writing and editing*
- *GC, CDF, GEM - primary investigator(s); conceived, coordinated, and directed all study activities, responsible for data collection, project integrity, manuscript content and editorial oversight and correspondence*

## Prior presentation

Preliminary analyses and portions of this data were accepted for a poster presentation at the 2019 American Society of Radiation Oncology (ASTRO) Annual Meeting, September 25–28, 2019, Chicago, IL, USA.

A detailed description of the spatial neighborhood similarity method used in this work has been published in the Journal of Biomedical Informatics in 2020.

A brief description of a subpart of Fig. 2 was accepted as a short conference presentation at the 2020 IEEE VIS conference, October 2020, Salt Lake City, UT, USA.

## Data sharing statement

Research data is not available at this time.

## Conflict of interest statement

The authors have nothing to disclose.

## Acknowledgements/funding disclosures

Dr. van Dijk receives funding from the Nederlandse Organisatie voor Wetenschappelijk Onderzoek/Netherlands Research Organization Rubicon Award (452182317).

Dr. Mohamed receives funding support from an MD Anderson Institutional Research Grant (IRG).

Dr.s Canahuate, Fuller, Mohamed, Vock and Marai received/receive funding and salary support related to this project during the period of study execution from: the National Institutes of Health (NIH) Big Data to Knowledge (BD2K) Program of the National Cancer Institute (NCI) Early Stage Development of Technologies in Biomedical Computing, Informatics, and Big Data Science Award (1R01CA2148250), and from the QuBBD program of the National Cancer Institute (R01CA225190); National Science Foundation (NSF), Division of Mathematical Sciences, Joint NIH/NSF Initiative on Quantitative Approaches to Biomedical Big Data (QuBBD) Grant (NSF 1557679).

Drs. Fuller and Mohamed received/receive funding and salary support related to this project during the period of study execution from the NCI Early Phase Clinical Trials in Imaging and Image-Guided Interventions Program (1R01CA218148); an NIH/NCI Cancer Center Support Grant (CCSG) Pilot Research Program Award from the UT MD Anderson CCSG Radiation Oncology and Cancer Imaging Program (P30CA016672).

Dr. Fuller and Mohamed received/receive funding and salary support unrelated to this project during the period of study execution from: the National Institute for Dental and Craniofacial Research Establishing Outcome Measures Award (1R01DE025248/R56DE025248) and Academic Industrial Partnership Grant (R01DE028290).

Dr. Fuller received/receives additional funding and salary support unrelated to this project during the period of study execution from: National Institute of Biomedical Imaging and Bioengineering (NIBIB) Research Education Programs for Residents and Clinical Fellows Grant (R25EB025787-01); an NIH/NCI Head and Neck Specialized Programs of Research Excellence (SPORE) Developmental Research Program Award (P50 CA097007); NSF Division of Civil, Mechanical, and Manufacturing Innovation (CMMI) grant (NSF 1933369); and the Sabin Family Foundation. Direct infrastructure support for this project is provided to Dr. Fuller by the multidisciplinary Stiefel Oropharyngeal Research Fund of the University of Texas MD Anderson Cancer Center Charles and Daneen Stiefel Center for Head and Neck Cancer and the Cancer Center Support Grant (P30CA016672) and the MD Anderson Program in Image-guided Cancer Therapy. Dr. Fuller has received direct industry grant support, honoraria, and travel funding from Elekta AB unrelated to this project.

Dr. Marai received/receives additional funding, travel funding, and salary support unrelated to this project during the period of study execution from the U.S. National Institutes of Health (NIH) National Library of Medicine (R01LM012527), from the U.S. National Science Foundation (NSF) through five other awards (CNS-1625941, CNS-1828265, CDS&E-1854815, IIS-2031095, CBET-1854815), from the Discovery Partners Institute in Chicago, and from The Joseph and Bessie Feinberg Foundation.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.radonc.2021.06.016>.

## References

- [1] de Carvalho AC, Perdomo S, dos Santos W, Fernandes GC, de Jesus LM, Carvalho RS, et al. Impact of genetic variants in clinical outcome of a cohort of patients with oropharyngeal squamous cell carcinoma. *Sci Rep* 2020;10:6.
- [2] Pytynia KB, Dahlstrom KR, Sturgis EM. Epidemiology of HPV-associated oropharyngeal cancer. *Oral Oncol* 2014;50:380–6.
- [3] Adelstein DJ, Ridge JA, Gillison ML, Chaturvedi AK, D'Souza G, Gravitt PE, Westra W, Psyrrri A, Martin Kast W, Koutsky LA, et al. Head and neck squamous cell cancer and the human papillomavirus: summary of a National Cancer Institute State of the Science Meeting, November 9–10, 2008, Washington, DC. *Head Neck* 2009;31:1393–422.
- [4] Massarelli E, Ferrarotto R, Glisson BS. New strategies in human papillomavirus-related oropharynx cancer: effecting advances in treatment for a growing epidemic. *Clin Cancer Res* 2015;21:3821–8.
- [5] You EL, Henry M, Zeitouni AG. Human papillomavirus-associated oropharyngeal cancer: review of current evidence and management. *Curr Oncol* 2019;26:119.
- [6] Young D, Xiao CC, Murphy B, Moore M, Fakhry C, Day TA. Increase in head and neck cancer in younger patients due to human papillomavirus (HPV). *Oral Oncol* 2015;51:727–30.
- [7] Marai GE, Ma C, Burks AT, Pellolio F, Canahuate G, Vock DM, et al. Precision risk analysis of cancer therapy with interactive nomograms and survival plots. *IEEE Trans Visual Comput Graphics* 2019;25:1732–45.
- [8] Tosado J, Zdilar L, Elhalawani H, Elgohari B, Vock DM, Marai GE, Fuller C, Mohamed ASR, Canahuate G. Clustering of Largely right-censored oropharyngeal head and neck cancer patients for discriminative groupings to improve outcome prediction. *Sci Rep* 2020;10:3811.
- [9] Langendijk JA, Doornaert P, Verdonck-de Leeuw IM, Leemans CR, Aaronson NK, Slotman BJ. Impact of late treatment-related toxicity on quality of life among patients with head and neck cancer treated with radiotherapy. *J Clin Oncol* 2008;26:3770–6.
- [10] Manikantan K, Khode S, Sayed SI, Roe J, Nutting CM, Rhys-Evans P, Harrington KJ, Kazi R. Dysphagia in head and neck cancer. *Cancer Treatment Rev* 2009;35:724–32.
- [11] Nguyen NP, Frank C, Moltz CC, Vos P, Smith HJ, Nguyen PD, Martinez T, Karlsson U, Dutta S, Lemanski C, Nguyen LM, Sallah S. Analysis of factors influencing aspiration risk following chemoradiation for oropharyngeal cancer. *Br J Radiol* 2009;82:675–80.
- [12] Hunter KU, Lee OE, Lyden TH, Haxer MJ, Feng FY, Schipper M, Worden F, Prince ME, McLean SA, Wolf GT, Bradford CR, Chepeha DB, Eisbruch A. Aspiration pneumonia after chemo-intensity-modulated radiation therapy of oropharyngeal carcinoma and its clinical and dysphagia-related predictors. *Head Neck* 2014;36:120–5.
- [13] Nguyen NP, Frank C, Moltz CC, Karlsson U, Nguyen PD, Ward HW, Vos P, Smith HJ, Huang S, Nguyen LM, Lemanski C, Ludin A, Sallah S. Analysis of factors influencing Dysphagia severity following treatment of head and neck cancer. *Anticancer Res* 2009;29:3299–304.
- [14] Setton J, Lee NY, Riaz N, Huang S-H, Waldron J, O'Sullivan B, Zhang Z, Shi W, Rosenthal DI, Hutcheson KA, Garden AS. A multi-institution pooled analysis of gastrostomy tube dependence in patients with oropharyngeal cancer treated with definitive intensity-modulated radiotherapy. *Cancer* 2015;121:294–301.
- [15] Wu J, Gensheimer MF, Zhang N, Han F, Liang R, Qian Y, Zhang C, Fischbein N, Pollom EL, Beadle B, Le Q-T, Li R. Integrating tumor and nodal imaging characteristics at baseline and mid-treatment computed tomography scans to predict distant metastasis in oropharyngeal cancer treated with concurrent chemoradiotherapy. *Int J Radiat Oncol Biol Phys* 2019;104:942–52.
- [16] Wentzel A, Hanula P, van Dijk LV, Elgohari B, Mohamed ASR, Cardenas CE, Fuller CD, Vock DM, Canahuate G, Marai GE. Precision toxicity correlates of tumor spatial proximity to organs at risk in cancer patients receiving intensity-modulated radiotherapy. *Radiother Oncol* 2020;148:245–51.
- [17] Luciani T, Wentzel A, Elgohari B, Elhalawani H, Mohamed A, Canahuate G, et al. A spatial neighborhood methodology for computing and analyzing lymph node carcinoma similarity in precision medicine. *J Biomed Inf: X* 2020;5:100067.
- [18] Grégoire V, Ang K, Budach W, Grau C, Hamoir M, Langendijk JA, Lee A, Le Q-T, Maingon P, Nutting C, O'Sullivan B, Porceddu SV, Lengele B. Delineation of the neck node levels for head and neck tumors: a 2013 update. DAHANCA, EORTC, HKNPCSG, NCIC CTG, NCRI, RTOG, TROG consensus guidelines. *Radiother Oncol* 2014;110:172–81.
- [19] U. D. of Health, Human Services and others, "Common terminology criteria for adverse events (CTCAE) version 4.0," National Institutes of Health, National Cancer Institute, 2009.
- [20] Elhalawani H, Lin TA, Volpe S, Mohamed ASR, White AL, Zafereo J, et al. Machine learning applications in head and neck radiation oncology: lessons from open-source radiomics challenges. *Front Oncol* 2018;8:294.
- [21] Elhalawani H, Mohamed ASR, White AL, Zafereo J, Wong AJ, Berends JE, AboHashem S, Williams B, Aymard JM, Kanwar A, Perni S, Rock CD, Cooksey L, Campbell S, Ding Y, Lai SY, Marai EG, Vock D, Canahuate GM, Freymann J, Farahani K, Kalpathy-Cramer J, Fuller CD. Matched computed tomography segmentation and demographic data for oropharyngeal cancer radiomics challenges. *Scientific Data* 2017;4.
- [22] Upton GJG. Fisher's exact test. *J Royal Statist Soc: A (Statistics in Society)* 1992.
- [23] McKinney W et al. Data structures for statistical computing in python. *Proceedings of the 9th Python in Science Conference*, 2010.
- [24] Wentzel A, van Dijk CL, Mohamed ASR, Fuller CD, Marai GE. Explainable spatial clustering, leveraging spatial data in radiation oncology, *IEEE Transactions on Visualization and Computer Graphics*, in press.
- [25] Marai GE, Pinaud B, Bühler K, Lex A, Morris JH. Ten simple rules to create biological network figures for communication., *PLoS Comput Biol*, no. 9, p. e1007244, 9 2019.
- [26] Christopherson KM, Ghosh A, Mohamed ASR, Kamal M, Gunn GB, Dale T, Kalpathy-Cramer J, Messer J, Garden AS, Elhalawani H, et al. Chronic radiation-associated dysphagia in oropharyngeal cancer survivors: towards age-adjusted dose constraints for deglutitive muscles. *Clin Transl Radiat Oncol* 2019;18:16–22.
- [27] Mehanna H, Robinson M, Hartley A, Kong A, Foran B, Fulton-Lieuw T, Dalby M, Mistry P, Sen M, O'Toole L, Al Booz H, Dyker K, Moleron R, Whitaker S, Brennan S, Cook A, Griffin M, Aynsley E, Rolles M, De Winton E, Chan A, Srinivasan D, Nixon I, Grumett J, Leemans CR, Buter J, Henderson J, Harrington K, McConkey C, Gray A, Dunn J. Radiotherapy plus cisplatin or cetuximab in low-risk human papillomavirus-positive oropharyngeal cancer (De-ESCaLATE HPV): an open-label randomised controlled phase 3 trial. *Lancet (London, England)* 2019;393:51–60.
- [28] Gillison ML, Trotti AM, Harris J, Eisbruch A, Harari PM, Adelstein DJ, Jordan RCK, Zhao W, Sturgis EM, Burtness B, Ridge JA, Ringash J, Galvin J, Yao M, Koyfman SA, Blakaj DM, Razaq MA, Colevas AD, Beitler JJ, Jones CU, Dunlap NE, Seaward SA, Spencer S, Galloway TJ, Phan J, Dignam JJ, Le QT. Radiotherapy plus cetuximab or cisplatin in human papillomavirus-positive oropharyngeal cancer (NRG Oncology RTOG 1016): a randomised, multicentre, non-inferiority trial. *Lancet (London, England)* 2019;393:40–50.
- [29] Wentzel A, Hanula P, Luciani T, Elgohari B, Elhalawani H, Canahuate G, et al. Cohort-based T-SSIM Visual Computing for Radiation Therapy Prediction and Exploration. *IEEE Trans Visual Comput Graphics* 2020;26:949–59.