European Radiology
https://doi.org/10.1007/s00330-019-06159-y

IMAGING INFORMATICS AND ARTIFICIAL INTELLIGENCE

Head and neck squamous cell carcinoma: prediction of cervical lymph node metastasis by dual-energy CT texture analysis with machine learning

Reza Forghani[1,2,3,4] · Avishek Chatterjee[5] · Caroline Reinhold[1,3] · Almudena Pérez-Lara[2,6] · Griselda Romero-Sanchez[2] · Yoshiko Ueno[3,7] · Maryam Bayat[2] · James W. M. Alexander[2] · Lynda Kadi[2,8] · Jeffrey Chankowsky[3] · Jan Seuntjens[4,5] · Behzad Forghani[1,4]

Received: 8 January 2019 / Revised: 27 February 2019 / Accepted: 13 March 2019

Abstract

Objectives This study was conducted in order to evaluate a novel risk stratification model using dual-energy CT (DECT) texture analysis of head and neck squamous cell carcinoma (HNSCC) with machine learning to (1) predict associated cervical lymphadenopathy and (2) compare the accuracy of spectral versus single-energy (65 keV) texture evaluation for endpoint prediction.

Methods Eighty-seven patients with HNSCC were evaluated. Texture feature extraction was performed on virtual monochromatic images (VMIs) at 65 keV alone or different sets of multi-energy VMIs ranging from 40 to 140 keV, in addition to iodine material decomposition maps and other clinical information. Random forests (RF) models were constructed for outcome prediction with internal cross-validation in addition to the use of separate randomly selected training (70%) and testing (30%) sets. Accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were determined for predicting positive versus negative nodal status in the neck.

Results Depending on the model used and subset of patients evaluated, an accuracy, sensitivity, specificity, PPV, and NPV of up to 88, 100, 67, 83, and 100%, respectively, could be achieved using multi-energy texture analysis. Texture evaluation of VMIs at 65 keV alone or in combination with only iodine maps had a much lower accuracy.

Conclusions Multi-energy DECT texture analysis of HNSCC is superior to texture analysis of 65 keV VMIs and iodine maps alone and can be used to predict cervical nodal metastases with relatively high accuracy, providing information not currently available by expert evaluation of the primary tumor alone.

Key Points

- *Texture features of HNSCC tumor are predictive of nodal status.*
- *Multi-energy texture analysis is superior to analysis of datasets at a single energy.*
- *Dual-energy CT texture analysis with machine learning can enhance noninvasive diagnostic tumor evaluation.*

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00330-019-06159-y) contains supplementary material, which is available to authorized users.

✉ Reza Forghani
reza.forghani@mcgill.ca

[1] Department of Radiology and Research Institute of the McGill University Health Centre, McGill University, Room C02.5821, 1001 Decarie Blvd, Montreal, QC H4A 3J1, Canada

[2] Segal Cancer Centre and Lady Davis Institute for Medical Research, Jewish General Hospital, Room C-212.1, 3755 Cote Ste-Catherine Road, Montreal, QC H3T 1E2, Canada

[3] Department of Radiology, Royal Victoria Hospital, McGill University Health Centre, 1001 Decarie Blvd, Montreal, QC H4A 3J1, Canada

[4] Gerald Bronfman Department of Oncology, McGill University, Montreal, QC, Canada

[5] Medical Physics Unit, Cedars Cancer Centre, McGill University Health Centre, 1001 Decarie Blvd, Montreal, QC H4A 3J1, Canada

[6] Department of Radiology, Hospital Regional Universitario de Málaga, Avenida Carlos Haya, S/N, 29010 Málaga, Spain

[7] Department of Radiology, Kobe University Graduate School of Medicine, 7-5-2, Kusunoki-cho, Chuo-ku, Kobe, Hyogo 650-0017, Japan

[8] Faculty of Medicine, Université de Montréal, Montreal, QC, Canada

Published online: 12 April 2019

Keywords Multidetector computed tomography · Machine learning · Artificial intelligence · Head and neck neoplasms · Computer-assisted diagnosis

Abbreviations

AUC	Area under the receiver operating curve
CI	Confidence intervals
DECT	Dual-energy CT
HNSCC	Head and neck squamous cell carcinoma
NPV	Negative predictive value
PPV	Positive predictive value
RF	Random forests
ROI	Region of interest
VMI	Virtual monochromatic image

Introduction

Accurate detection of metastatic cervical lymphadenopathy is essential for optimal staging and management of patients with head and neck squamous cell carcinoma (HNSCC). Despite significant improvements with current imaging techniques, detection of early nodal metastases remains a challenge [1–3]. As a result, HNSCC patients with a clinically N0 neck harboring tumors at high-risk sites with greater than 15–20% risk of associated nodal metastases routinely undergo neck dissections, even if this could result in overtreatment of potentially 60–70% of patients [4–7]. Because of the high percentage of patients potentially undergoing an unnecessary invasive procedure with potential for significant associated complications and morbidity, there is significant interest in predictive models that may increase diagnostic accuracy, including negative predictive value, sufficiently to reduce unnecessary elective neck dissections [6, 8].

Texture or radiomic analysis has been used in different organ systems to predict various endpoints of interest including tumor molecular features, patient prognosis, and response to treatment [9–16]. A few studies have shown potentially promising results in applying texture analysis for the evaluation of pathologic nodes in patients with lung cancer [17, 18]. However, there can be significant barriers to such an approach in the clinically N0 neck where by default, there is no obvious target for analysis and successful characterization of all relevant small nodes could be challenging. On the other hand, an alternative approach may be to use the texture features of the tumor itself to predict the risk of nodal metastases, similar to texture analysis of soft tissue sarcomas to predict the risk of associated lung metastases or application of texture in prognostic models [9, 11, 15, 16].

So far, texture studies have included analysis of conventional (single-energy) CT, MRI, and PET, but there are few texture studies using different dual-energy CT (DECT) image sets [19–22], where the additional quantitative spectral information on tissue attenuation [23–29] could potentially improve predictive models. It has been shown that HNSCC and normal tissues have different energy-dependent attenuation characteristics, i.e., different attenuation curves at different virtual monochromatic image (VMI) energies, and that low energy VMIs improve HNSCC visualization [23–27]. Analysis of different energy VMI curves has also been used for lymph node characterization [30]. Furthermore, iodine maps generated from DECT scans have been used for evaluation of HNSCC or cervical lymph nodes [28, 29, 31]. One recent study evaluating benign parotid tumors also suggests that texture analysis of multi-energy datasets can improve predictive model performance [22], but it is not yet known whether those observations can be more widely applied to other tumors and endpoints.

In this study, we (1) evaluated the feasibility of using machine learning methods to analyze texture features extracted from DECT scans of HNSCC patients to predict the presence of cervical metastatic lymphadenopathy based on the intrinsic features of the index tumor and (2) compared the performance of analysis using multi-energy VMI datasets and DECT-derived iodine maps to analysis of VMIs at a single energy (65 keV), typically considered similar to a standard single-energy CT acquisition [24, 32–35].

Materials and methods

Patient population

Institutional review board approval and waiver of informed consent were obtained for this single institution retrospective study. A database search revealed 567 scans from patients with a primary head and neck malignancy (excluding lymphoma or leukemia) acquired as a DECT scan between June 2013 and August 2016. After application of inclusion and exclusion criteria (listed in Table 1), the final study population consisted of 87 patients with histopathology-proven HNSCC, consisting of a mix of recurrent and untreated tumors (Fig. 1). For this pilot feasibility study, nodal status was determined based on final stage determined at the tumor board for the larger population, with postoperative gold standard pathologic confirmation of the nodal status in a subset of patients (Fig. 1). The latter were used exclusively for evaluating testing (prediction) model performance in the subset of untreated patients (Fig. 1).

Dual-energy CT imaging, image postprocessing, and multi-energy DECT approach and data preparation for texture analysis

Dual-energy CT is performed as part of standard of care for head and neck cancer patients at our institution. All

Table 1 Inclusion and exclusion criteria for patient selection

Inclusion criteria	Exclusion criteria
Patient with pathology proven HNSCC who had undergone a scan at the time of initial tumor workup or identification of recurrence	Nasopharyngeal primary Cervical metastatic disease without an identifiable primary tumor Unavailable or ambiguous neck dissection or tumor board nodal stage Tumor that is too small (less than 5 mm maximal diameter on axial images) or not clearly visualized on at least one CT slice Entire tumor volume obscured by dental or severe motion artifact

Baseline characteristics of patient population and segmentation information

Total population = 87; 64 untreated and 23 recurrent tumors
Average age 69 years (range 43 to 96)
32 women, 55 men
44 patients had positive cervical nodal status, 43 had negative cervical nodal status
Segmentation:
Average ROI size (±SD) was 3.7 ± 3.4 cm^2 for untreated tumors and 7.5 ± 12.2 cm^2 for recurrent tumors

scans were obtained with the same 64-slice dual-energy scanner (GE Discovery CT750HD; GE Healthcare), using the same technique after administration of 80 mL of iopamidol (Isovue 300®, Bracco Diagnostics Inc.), injected at a rate of 2 mL/s, and the patient was scanned after a delay of 65 s as described in greater detail in the Electronic supplementary material or previously described [24, 26]. To take advantage of the large amount of quantitative spectral data generated using DECT, multi-energy VMIs were reconstructed from 40 to 140 keV in 5 keV increments at the GE Advantage workstation (4.6; GE Healthcare). This was modeled to capture the energy-dependent attenuation changes of the tumor (example in Fig. 2) [22–24, 26, 36]. This resulted in 21 different reconstructions per case, or a total of 1827 image sets for the total group of 87 patients that were used for multi-

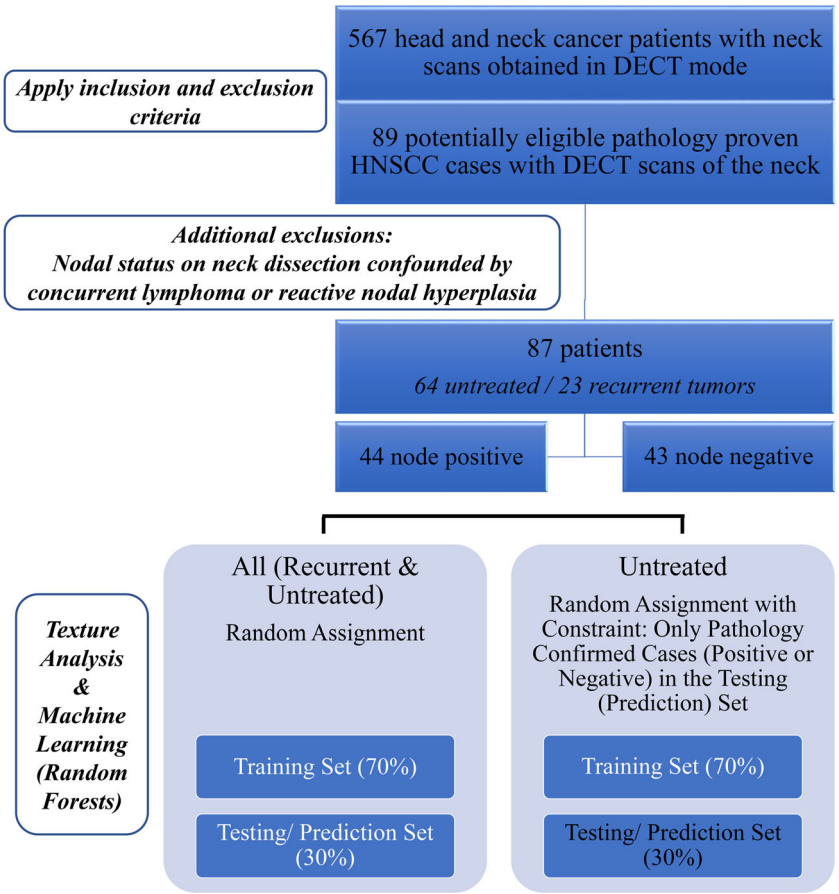

Fig. 1 Flowchart of the study population and assignments for texture analysis and machine learning for construction of prediction models

energy texture analysis. In addition, iodine–water material decomposition maps were also created, reflecting the estimated concentration and distribution of iodine within different tissues and tumor [28, 29, 36–38] (Fig. 2). These images were generated both in gray-scale and color for this analysis.

Tumor segmentation, texture feature extraction, and other features evaluated

Texture analysis of HNSCC was performed using a (1) commercially available software for different energy VMIs (TexRAD Ltd) or (2) an in-house developed program for analysis of iodine maps (see below). For each case, an ROI was manually drawn around the tumor on its largest cross-sectional area [39] on the image set where the tumor was best seen, typically the 40-keV VMIs [24, 33], in conjunction with VMIs at other energies if needed. ROIs were either drawn (1) by a head and neck radiologist with over 6-years postfellowship experience (R.F.) or (2) first drawn by a head and neck radiology fellow with at least 1 year of neuroradiology fellowship experience and then reviewed by the same attending head and neck radiologist (with revisions if necessary). The ROI was then used for analysis and texture feature extraction from different combinations of different energy VMIs and iodine maps, or the 65-keV VMIs only; 65 keV VMIs were selected because either 65 or 70 keV are the VMI energies that are typically considered equivalent and used as a replacement for a conventional 120 kVp single-energy neck CT acquisition when a scan is acquired in DECT mode [24, 32–35].

Following tumor segmentation, six texture features based on first-order statistics of the gray level density histogram, each analyzed using different in-plane filtration settings, were derived from each image set using TexRAD software for all VMI energy datasets [39–41] (Table 2). For texture analysis of iodine maps, an in-house program was developed using MATLAB. This is because this type of image could not be analyzed either with TexRAD or 3DSlicer, a freely available segmentation and processing tool (https://www.slicer.org/). Using the in-house developed program, the first-order statistical features corresponding to features implemented in TexRAD were computed (Table 2). Figure 3 provides a graphical example of changes in extracted texture features at different VMI energies. The segmentation criteria and texture feature extraction process, including the use of in-house program, are described in greater detail in the Electronic supplementary material.

For constructing prediction models using machine learning (see below), analysis was first performed using texture features extracted from the 21 VMI energy set and compared to analysis of texture features from a dataset at a single energy of 65 keV, as discussed above. Thereafter, for the main subset of untreated patients with exclusively histopathology-proven positive or negative nodal metastases (Fig. 1), multiple additional analyses were performed. We tested a model using a smaller set of 11 VMIs at intervals of 10 keV to see whether a reduced number of VMIs within the same range could be used to achieve a similar performance. We also tested texture features extracted from iodine maps. Lastly, models were tested using texture features in combination with basic patient information (patient's sex) or tumor size in three dimensions (transverse, anteroposterior, craniocaudal, or maximal tumor size regardless of plane; each tested independently as a potential predictor). This was done to evaluate for any potential contributory effect vis-a-vis the texture features.

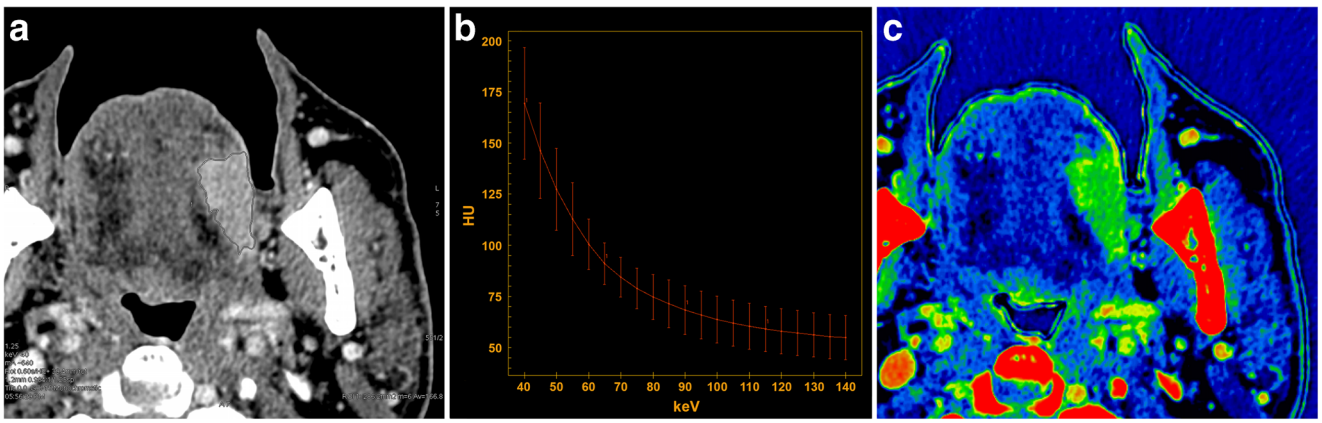

Fig. 2 Example of energy-dependent attenuation changes of HNSCC and iodine map. **a** Region of interest analysis of a left oral tongue cancer on a low-energy virtual monochromatic image (VMI) at 40 keV with the corresponding (**b**) spectral Hounsfield unit attenuation curve and (**c**) iodine material decomposition map are shown. The curve demonstrates tumor attenuation at 21 different VMI energy levels, ranging between 40 and 140 keV in increments of 5 keV. In this study, VMIs reconstructed at 21 different energies were used for multi-energy texture analysis of tumors. As a comparison, an acquisition from a conventional single-energy CT scan would result in a single measurement set, typically corresponding to what is seen on a 65- or 70-keV VMI. (**c**) Alternatively, DECT acquisitions can be used to generate material basis decomposition maps, such as the iodine map reflecting the estimated concentration and distribution of iodine within a tumor. This type of map was also used for texture analysis in this study

Table 2 Texture features analyzed

First-order statistics	In-plane filtration step (applied to each of the first-order feature sets)
(1) Average density values (2) Standard deviation (SD) (3) Mean of positive pixels (4) Entropy (5) Skewness (6) Kurtosis	Laplacian of Gaussian spatial band-pass filter at different spatial scale filter (SSF) values: (1) SSF 0: no filter (2–6) SSF 2–6: values between 2 and 6 to enhance and extract CT attenuation fine-texture features with filters with a radius ranging between 2 and 6 mm

Machine learning and statistical analysis

The random forests (RF) method [42, 43] was used to build the prediction models for the different outcomes. To build the most accurate prediction model for each outcome, a computer program was written that generated and evaluated hundreds of models. For an unbiased assessment of the model accuracy [44], 30% of the patients were randomly selected and set aside as the test (prediction) group. The remaining 70% were used to train and validate each prediction model during the search. For the validation, i.e., the training accuracy, the internal check in RF itself was used. During the model selection process, the final hyperparameters that were selected were those that resulted in the best prediction accuracy and a compromise between being underfit and overfit. Using the prediction models, the outcome for the test cases was predicted and the accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and area under the receiver operating curve (AUC) of the test cases were determined (prediction accuracy). The R package (R Development Core Team (2008), ISBN 3-900051-07-0, http://www.R-project.org) was used for the machine learning and statistical analysis. The R package, randomForest, was used for building the random

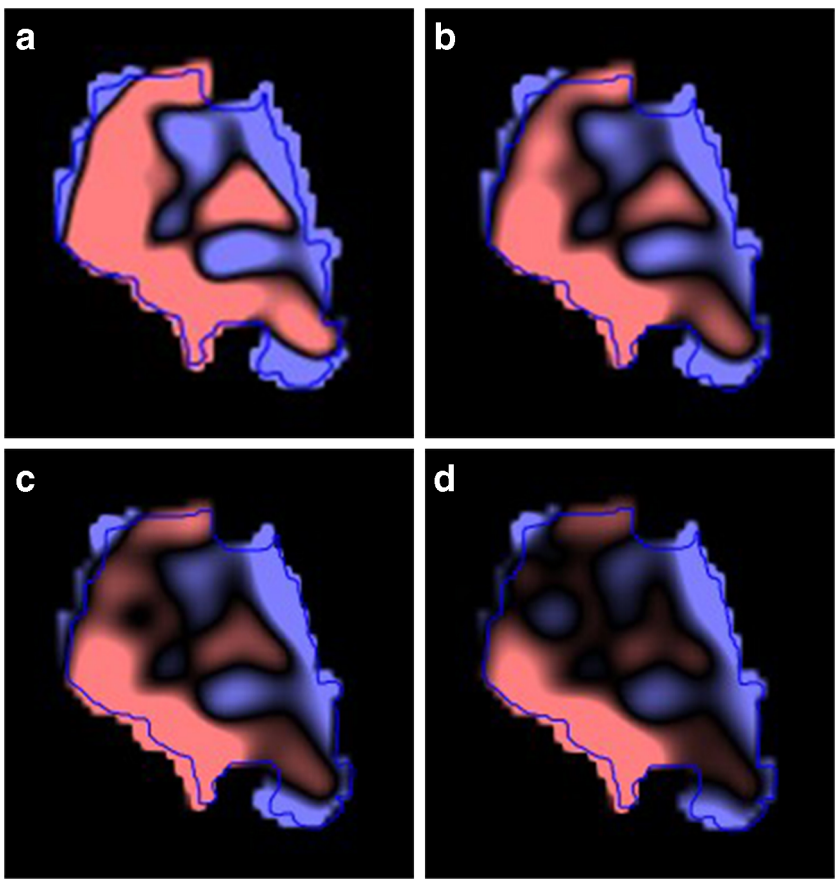

Fig. 3 Graphical example of changes in extracted texture maps at different virtual monochromatic image (VMI) energies. Selected examples consisting of texture maps extracted from (**a**) 40 keV, (**b**) 65 keV, (**c**) 80 keV, and (**d**) 140 keV VMIs from a supraglottic cancer are shown. Note the changes in texture maps at different energies. The additional information and changes seen on different energy VMIs are a reflection of tissue-specific characteristics and likely underlie the improvements in prediction models seen using multi-energy VMIs compared to VMIs at a single energy

Table 3 HNSCC tumor site and stage (all patients)

Primary site	Number of patients (total 87)	T stage	N stage	M stage
Oral cavity	31 (23 U/8 R)	T1 (9)	N0 (8)	M0 (23)
		T2 (4)	N1 (3)	M1 (0)
		T3 (4)	N2 (11)	
		T4 (6)	N3 (1)	
		R (8)	R (8)	R (8)
Oropharynx	22 (18 U/4 R)	T1 (6)	N0 (6)	M0 (17)
		T2 (6)	N1 (9)	M1 (1)
		T3 (2)	N2 (3)	
		T4 (4)	N3 (0)	
		R (4)	R (4)	R (4)
Hypopharynx or larynx	24 (18 U/6 R)	T1 (1)	N0 (13)	M0 (18)
		T2 (2)	N1 (0)	M1 (0)
		T3 (11)	N2 (4)	
		T4 (4)	N3 (1)	
		R (6)	R (6)	R (6)
Other	10 (5 U/5 R)	T1 (0)	N1 (0)	M0 (5)
		T2 (2)	N2 (0)	M1 (0)
		T3 (1)	N3 (0)	
		T4 (2)		
		R (5)	R (5)	R (5)

Table summarizing the characteristics of the entire set of patients evaluated in this study. In this population, cervical nodal status was based on the final stage determined at the tumor board or postoperative pathologic confirmation of nodal status for the subset of patients that underwent surgery and lymph node dissection

U, untreated; *R*, recurrent

forests models [45]. For a more detailed description of the machine learning approach and model construction, please refer to the Electronic supplementary material.

Results

Patient demographics and HNSCC sites

Eighty-seven patients met the inclusion criteria for this study. For additional details on the distribution of the primary tumor sites and baseline characteristics, please refer to Fig. 1 or Tables 1, 3, and E1 (Electronic supplementary material).

Comparison of prediction models for all patients versus untreated subset using multi-energy and single-energy datasets

We first derived and compared prediction models based on the entire population (untreated and recurrent cases) and the more homogeneous untreated subset for determination of the presence of cervical lymphadenopathy. Using texture analysis of multi-energy datasets consisting of 21 VMIs, the independent testing (prediction) models had an accuracy, sensitivity, specificity, PPV, and NPV of 76 (95% CI 59, 93), 77, 75, 77, and 75%, respectively, in the entire population consisting of the untreated and recurrent group and 88 (95% CI 65, 100), 100, 67, 83, and 100%, respectively, in the more homogeneous untreated group. When the 65-keV dataset was used in isolation, the prediction accuracy dropped substantially with accuracy, sensitivity, specificity, PPV, and NPV of 60 (95% CI 41, 79), 69, 50, 60, and 60% (entire population) and 63 (95% CI 41, 85), 70, 56, 64, and 63% (untreated group).

Multiparametric evaluation of untreated patients for predicting cervical nodal metastases

After the initial analysis demonstrating superior performance of multi-energy VMI texture analysis, we focused on the more homogeneous population of untreated tumors ($n = 64$) for additional analyses. In addition to using the 21 VMI energy datasets at intervals of 5 keV described above (accuracy 88%; 95% CI 65, 100), we tested the addition of iodine map texture analysis, patient's sex, and multidimensional tumor size, without any impact on model performance (accuracy 88%; 95% CI 65, 100; Table 4). We then evaluated a smaller set of 11 VMIs at intervals of 10 keV, to see whether a reduced number of VMIs within the

Table 4 Prediction accuracy for classification of cervical nodal metastasis using different approaches in the untreated group ($n = 64$) with exclusively pathologically proven positive or negative nodal metastasis in the testing (prediction) sets

Model no.	Model description	Accuracy	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
1	Multi-energy VMIs (21 set; 5 keV intervals)	88% (65, 100)	100	67	83	100
2	Multi-energy VMIs (21 set; 5 keV intervals) + iodine maps	88% (65, 100)	100	67	83	100
3	Multi-energy VMIs (21 set; 5 keV intervals) + patient sex and multidimensional tumor measurements	88% (65, 100)	100	67	83	100
4	Multi-energy VMIs (11 set; 10 keV intervals)	88% (65, 100)	100	67	83	100
5	65 keV VMIs + iodine maps	63% (41, 85)	70	56	64	62

Results shown are for the independent testing (prediction) set. Please refer to the manuscript text for additional details and explanations

same range could be used to achieve a similar performance. There was only a marginal effect on overall model performance, with an identical accuracy (accuracy 88%; 95% CI 65, 100; Table 4). We also tested the combination of texture analysis on 65 keV VMIs and iodine maps to evaluate whether texture analysis of iodine maps can be used as a substitute for the multi-energy VMIs. However, performance dropped substantially, with an accuracy of 63% (95% CI 41, 85; Table 4). Six classifiers were used in all final models except for the 11 VMI set model in which 9 classifiers were used (VMI range 40 and 130 keV; Table 5). Figure 4 demonstrates an example of a node that was negative by imaging but positive on pathology and correctly predicted using texture analysis of the primary tumor with machine learning.

Table 5 HNSCC tumor evaluation for prediction of cervical nodal metastases: final classifiers used in various prediction models for untreated HNSCC patients

Models 1–3	Model 4
SSF0.sd.75	SSF0.sd.70
SSF2.entropy.55	SSF2.entropy.50
SSF0.mpp.125	SSF3.mpp.50
SSF2.kurtosis.40	SSF0.mean.130
SSF4.mpp.40	SSF5.mean.50
SSF0.entropy.80	SSF2.kurtosis.40
	SSF0.entropy.80
	SSF5.skewness.80
	SSF0.kurtosis.130

Model numbers are based on the preceding table (Table 4). The final classifiers were identical for models 1–3 and therefore shown in one column. The order shown reflects classifier importance (top—most important). Classifier designations represented as *spatial scale filter.texture features based on first-order statistics of the gray level density histogram.VMI energy*. Potential spatial scale filters were 0, 2, 3, 4, 5, or 6. Potential texture features consisted of (1) average density values (mean), (2) standard deviation (sd), (3) mean of positive pixels (mpp), (4) entropy, (5) skewness, and (6) kurtosis. Potential VMI energies ranged between 40 and 140 keV in 5 keV increments for multi-energy analysis or 65 keV only for single-energy analysis. Please refer to the text for additional details

Discussion

In this feasibility study, we demonstrate that risk stratification and prediction of neck nodal status using texture features of the index tumor is feasible and machine learning using texture analysis of multi-energy DECT datasets is superior to analysis of single-energy datasets at 65 keV. A radiomic model combined with machine learning can be performed on studies already obtained as part of the patient's initial workup, is noninvasive, and can be performed preoperatively. These are advantages compared to other approaches under consideration. For example, there are investigations attempting to use advanced molecular analysis of the resected tumor for risk stratification [8, 46]. However, even if molecular markers with sufficient accuracy are eventually identified, these by default require robust molecular analytic platforms that can produce results rapidly enough for the surgeon to await those results following tumor resection in order to make a determination of whether or not to proceed with neck dissection. These represent significant practical and structural barriers. Other approaches, such as those relying on sentinel node biopsy, still entail additional invasive components, unlike the radiomic approach. Figure 4 demonstrates an example of correct prediction of micrometastases in case of imaging negative nodes. However, equally if not more important is the high NPV and concordance with imaging negative nodes. If validated in larger studies, this approach could be combined with radiologist's expert evaluation of nodes as a clinical assistant tool, potentially increasing diagnostic confidence in prediction of absent nodal metastases and reducing negative neck dissections.

In addition to the above, our investigation also demonstrates a potential additional value of spectral data available from DECT scans for radiomic analysis. In a recent study of benign parotid tumors, multi-energy analysis was shown to be superior to the analysis of datasets at a single-energy level (65 keV) for predicting tumor pathology. In this study, we also demonstrate superiority of multi-energy texture using a completely different, malignant tumor and a complex endpoint. Texture analysis performed on multi-energy datasets improved the accuracy of the prediction models compared to that performed solely on VMIs at 65 keV, typically considered

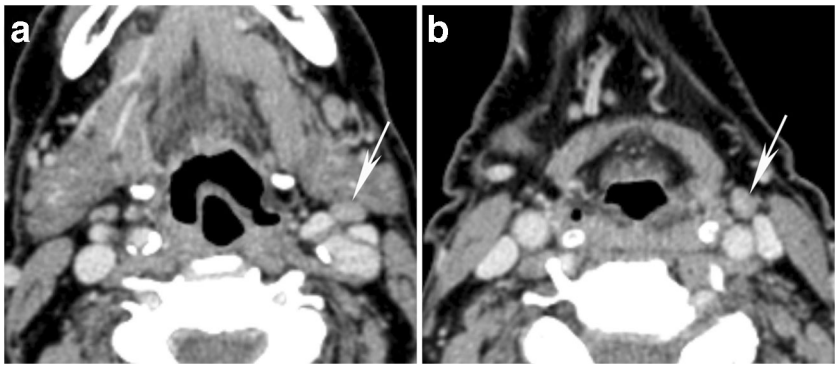

Fig. 4 Primary HNSCC tumor texture analysis and machine learning for prediction of associated nodal metastasis. Example of correct classification/prediction in a case of otherwise imaging negative, pathologically confirmed, nodal micrometastasis in this patient who underwent bilateral neck dissections. In this case, metastases were confirmed in resected specimens in two lymph nodes on the left (arrows), (**a**) one at left level II and the other (**b**) left level III. These nodes did not satisfy imaging criteria for abnormal nodes prospectively or retrospectively. The short axis diameter for each was 5 mm. On histopathology, the maximal micrometastatic focus measured 4 mm

equivalent to a standard 120-kVp single-energy CT acquisition [24, 32–35]. While the equivalence of texture features of 65 keV VMIs and single-energy CT requires independent investigations and validation, these results have potential implications beyond the specific application investigated in this study and could provide the foundation for application of the multi-energy radiomic approach to other pathology in the neck or in other organ systems.

The performance of an abbreviated 11 VMI set was almost the same as the 21 VMI set and suggests that this could be used as an alternative, which would be much less onerous from the standpoint of data processing and analysis. More classifiers were required using this approach (Table 5) compared to 21 VMI set analysis, but this may be a reflection of the small patient numbers used in this study and may improve if a larger set of patients were used. However, classifiers were selected from a wide range of energies, ranging from 40 keV VMIs that accentuate enhancing tissues the most to 130 keV VMIs that come close to simulating an unenhanced acquisition. Therefore, performing texture analysis on a wide range of VMIs appears important based on our results, similar to previous observations on parotid tumors [22]. Surprisingly, the iodine maps did not improve performance, suggesting that they cannot be used as replacement for the multi-energy datasets. From a theoretical standpoint, however, this is not a complete surprise since the multi-energy datasets capture the energy-dependent attenuation changes of different tissues and would be expected to have more information than any single reconstruction alone. The multi-energy radiomic approach could also be of interest for evaluating tumors in areas obscured by dental artifact, and these are topics of interest for future research using this approach.

This proof of concept study has many limitations. The main limitation of our study is the relatively small sample size analyzed for the determination of a complex risk stratification endpoint. Nonetheless, as a proof-of-principle study, all the models evaluated in this study support the feasibility of this approach, although requiring confirmation in larger and ideally prospective studies. We also did not evaluate the impact of contouring variation on the predictive models, although for comparison of single- versus multi-energy VMIs, the use of identical contours is necessary in order to eliminate other variables that may affect model performance. It would be mathematically incorrect to average many of the texture features (other than the mean) from more than one contour for the same tumor, but future studies evaluating the impact of contour variations on model performance would be of interest. Alternatively, in the future, autosegmentation approaches or direct image analysis using convolutional neural networks (for datasets that are sufficiently large) could be used that reduce or eliminate interrater variation and also enable a more workflow-friendly implementation. The confidence intervals are relatively large but this is not unexpected given the small sample size used. Even then, we took specific steps to avoid overfitting, including independent training and testing sets in addition to built-in mechanisms within RF itself. While it is likely that our observations can be applied to single-energy CT scans, this would require independent validation.

In conclusion, the novel proposed approach of using multi-energy texture analysis of the index HNSCC tumor can predict the cervical nodal status with relatively high accuracy, with a potential to complement the evaluation by expert radiologists. The high NPV in this preliminary study is particularly encouraging. If refined and validated in larger patient sets and possibly in combination with other imaging or clinical parameters, this approach has the potential to enable risk stratification that may eventually reach sufficient accuracy to reduce unnecessary neck dissections.

Funding This work was partly supported by a grant from the Rossy Cancer Network. R.F. is a clinical research scholar supported by the FRQS (Fonds de recherche en santé du Québec).

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is R. Forghani.

Conflict of interest The authors of this manuscript declare relationships with the following companies: R.F. has acted as consultant and speaker for GE Healthcare and is a founding partner and stockholder of 4Intel Inc. B.F. is a founding partner and stockholder of 4Intel Inc.

Statistics and biometry B.F. has significant statistical and informatics expertise and performed the mathematical and statistical analyses.

Informed consent Ethics approval was obtained by the Institutional Review Board of the Jewish General Hospital (CIUSSS West-Central Montreal).

Ethical approval Written informed consent was waived by the Institutional Review Board.

Methodology
- retrospective
- experimental
- performed at one institution

References

1. Som PM, Brandwein-Gensler MS (2011) Lymph nodes of the neck. In: Som PM, Curtin HD (eds) Head and neck imaging. Mosby, St. Louis
2. Forghani R, Yu E, Levental M, Som PM, Curtin HD (2014) Imaging evaluation of lymphadenopathy and patterns of lymph node spread in head and neck cancer. Expert Rev Anticancer Ther. https://doi.org/10.1586/14737140.2015.978862:1-18
3. Kostakoglu L (2011) PET/CT Imaging. In: Som PM, Curtin HD (eds) Head and neck imaging. Mosby, St. Louis
4. D'Cruz AK, Vaish R, Kapre N et al (2015) Elective versus therapeutic neck dissection in node-negative oral cancer. N Engl J Med 373:521–529
5. Abu-Ghanem S, Yehuda M, Carmel NN et al (2016) Elective neck dissection vs observation in early-stage squamous cell carcinoma of the oral tongue with no clinically apparent lymph node metastasis in the neck: a systematic review and meta-analysis. JAMA Otolaryngol Head Neck Surg 142:857–865
6. Paleri V, Urbano TG, Mehanna H et al (2016) Management of neck metastases in head and neck cancer: United Kingdom National Multidisciplinary Guidelines. J Laryngol Otol 130:S161–S169
7. Liao LJ, Hsu WL, Wang CT, Lo WC, Lai MS (2016) Analysis of sentinel node biopsy combined with other diagnostic tools in staging cN0 head and neck cancer: a diagnostic meta-analysis. Head Neck 38:628–634
8. Medina JE (2017) Cancer of the neck. In: Myers J, Hanna E, Myers EN (eds) Cancer of the head and neck. Wolters Kluwer, Philadelphia, pp 427–453
9. Zhang H, Graham CM, Elci O et al (2013) Locally advanced squamous cell carcinoma of the head and neck: CT texture and histogram analysis allow independent prediction of overall survival in patients treated with induction chemotherapy. Radiology 269:801–809
10. Aerts HJ, Velazquez ER, Leijenaar RT et al (2014) Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 5:4006
11. Parmar C, Grossmann P, Bussink J, Lambin P, Aerts HJ (2015) Machine learning methods for quantitative radiomic biomarkers. Sci Rep 5:13087
12. Dang M, Lysack JT, Wu T et al (2015) MRI texture analysis predicts p53 status in head and neck squamous cell carcinoma. AJNR Am J Neuroradiol 36:166–170
13. Buch K, Fujita A, Li B, Kawashima Y, Qureshi MM, Sakai O (2015) Using texture analysis to determine human papillomavirus status of oropharyngeal squamous cell carcinomas on CT. AJNR Am J Neuroradiol 36:1343–1348
14. Leijenaar RT, Carvalho S, Hoebers FJ et al (2015) External validation of a prognostic CT-based radiomic signature in oropharyngeal squamous cell carcinoma. Acta Oncol 54:1423–1429
15. Liu J, Mao Y, Li Z et al (2016) Use of texture analysis based on contrast-enhanced MRI to predict treatment response to chemoradiotherapy in nasopharyngeal carcinoma. J Magn Reson Imaging 44:445–455
16. Vallieres M, Freeman CR, Skamene SR, El Naqa I (2015) A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities. Phys Med Biol 60:5471–5496
17. Bayanati H, Thornhill RE, Souza CA et al (2015) Quantitative CT texture and shape analysis: can it differentiate benign and malignant mediastinal lymph nodes in patients with primary lung cancer? Eur Radiol 25:480–487
18. Andersen MB, Harders SW, Ganeshan B, Thygesen J, Torp Madsen HH, Rasmussen F (2016) CT texture analysis can help differentiate between malignant and benign lymph nodes in the mediastinum in patients suspected for lung cancer. Acta Radiol 57:669–676
19. Foncubierta-Rodriguez A, Jimenez del Toro OA, Platon A, Poletti PA, Muller H, Depeursinge A (2013) Benefits of texture analysis of dual energy CT for computer-aided pulmonary embolism detection. Conf Proc IEEE Eng Med Biol Soc 2013:3973–3976
20. Oldan J, He M, Wu T et al (2014) Pilot study: evaluation of dual-energy computed tomography measurement strategies for positron emission tomography correlation in pancreatic adenocarcinoma. J Digit Imaging 27:824–832
21. Depeursinge A, Foncubierta-Rodriguez A, Vargas A et al (2013) Rotation-covariant texture analysis of 4D dual-energy CT as an indicator of local pulmonary perfusion. 2013 IEEE 10th international symposium on biomedical imaging (ISBI), San Francisco, CA, USA, pp 145–148
22. Al Ajmi E, Forghani B, Reinhold C, Bayat M, Forghani R (2018) Spectral multi-energy CT texture analysis with machine learning for tissue classification: an investigation using classification of benign parotid tumours as a testing paradigm. Eur Radiol 28:2604–2611
23. Srinivasan A, Parker RA, Manjunathan A, Ibrahim M, Shah GV, Mukherji SK (2013) Differentiation of benign and malignant neck pathologies: preliminary experience using spectral computed tomography. J Comput Assist Tomogr 37:666–672
24. Lam S, Gupta R, Levental M, Yu E, Curtin HD, Forghani R (2015) Optimal virtual monochromatic images for evaluation of normal tissues and head and neck cancer using dual-energy CT. AJNR Am J Neuroradiol 36:1518–1524
25. Wichmann JL, Noske EM, Kraft J et al (2014) Virtual monoenergetic dual-energy computed tomography: optimization of kiloelectron volt settings in head and neck cancer. Invest Radiol 49:735–741
26. Forghani R, Levental M, Gupta R, Lam S, Dadfar N, Curtin HD (2015) Different spectral hounsfield unit curve and high-energy virtual monochromatic image characteristics of squamous cell carcinoma compared with nonossified thyroid cartilage. AJNR Am J Neuroradiol 36:1194–1200
27. Albrecht MH, Scholtz JE, Kraft J et al (2015) Assessment of an advanced monoenergetic reconstruction technique in dual-energy

computed tomography of head and neck cancer. Eur Radiol 25: 2493–2501
28. Tawfik AM, Razek AA, Kerl JM, Nour-Eldin NE, Bauer R, Vogl TJ (2014) Comparison of dual-energy CT-derived iodine content and iodine overlay of normal, inflammatory and metastatic squamous cell carcinoma cervical lymph nodes. Eur Radiol 24:574–580
29. Rizzo S, Radice D, Femia M et al (2018) Metastatic and non-metastatic lymph nodes: quantification and different distribution of iodine uptake assessed by dual-energy CT. Eur Radiol 28:760–769
30. Yang L, Luo D, Li L et al (2016) Differentiation of malignant cervical lymphadenopathy by dual-energy CT: a preliminary analysis. Sci Rep 6:31020
31. Yamauchi H, Buehler M, Goodsitt MM, Keshavarzi N, Srinivasan A (2016) Dual-energy CT-based differentiation of benign posttreatment changes from primary or recurrent malignancy of the head and neck: comparison of spectral Hounsfield units at 40 and 70 keV and iodine concentration. AJR Am J Roentgenol 206:580–587
32. Matsumoto K, Jinzaki M, Tanami Y, Ueno A, Yamada M, Kuribayashi S (2011) Virtual monochromatic spectral imaging with fast kilovoltage switching: improved image quality as compared with that obtained with conventional 120-kVp CT. Radiology 259:257–262
33. Forghani R, Kelly H, Yu E et al (2017) Low-energy virtual monochromatic dual-energy computed tomography images for the evaluation of head and neck squamous cell carcinoma: a study of tumor visibility compared with single-energy computed tomography and user acceptance. J Comput Assist Tomogr 41:565–571
34. Forghani R (2015) Advanced dual-energy CT for head and neck cancer imaging. Expert Rev Anticancer Ther. https://doi.org/10.1586/14737140.2015.1108193:1-13
35. Lam S, Gupta R, Kelly H, Curtin HD, Forghani R (2015) Multiparametric evaluation of head and neck squamous cell carcinoma using a single-source dual-energy CT with fast kVp switching: state of the art. Cancers (Basel) 7:2201–2216
36. Forghani R, Srinivasan A, Forghani B (2017) Advanced tissue characterization and texture analysis using dual-energy computed tomography: horizons and emerging applications. Neuroimaging Clin N Am 27:533–546
37. Forghani R, De Man B, Gupta R (2017) Dual-energy computed tomography: physical principles, approaches to scanning, usage, and implementation: part 2. Neuroimaging Clin N Am 27:385–400
38. Kraft M, Ibrahim M, Spector M, Forghani R, Srinivasan A (2018) Comparison of virtual monochromatic series, iodine overlay maps, and single energy CT equivalent images in head and neck cancer conspicuity. Clin Imaging 48:26–31
39. Ueno Y, Forghani B, Forghani R et al (2017) Endometrial carcinoma: MR imaging-based texture model for preoperative risk stratification—a preliminary analysis. Radiology. https://doi.org/10.1148/radiol.2017161950:161950
40. Parikh J, Selmi M, Charles-Edwards G et al (2014) Changes in primary breast cancer heterogeneity may augment midtreatment MR imaging assessment of response to neoadjuvant chemotherapy. Radiology 272:100–112
41. De Cecco CN, Ganeshan B, Ciolina M et al (2015) Texture analysis as imaging biomarker of tumoral response to neoadjuvant chemoradiotherapy in rectal cancer patients studied with 3-T magnetic resonance. Invest Radiol 50:239–245
42. Breiman L (2001) Random forests. Mach Learn 45:5–32
43. Caruana R, Karampatziakis N, Yessenalina A (2008) An empirical evaluation of supervised learning in high dimensions. ICML '08, Proceedings of the 25th International conference on Machine learning. ACM, Helsinki, pp 96–103
44. Hastie T, Tibshirani R, Friedman J (2009) The elements of statistical learning: Data mining, inference, and prediction, Second edition. Springer Series in Statistics, Springer-Verlag
45. Liaw A, Wiener M (2002) Classification and regression by randomForest. R News 2:18–22
46. Yang X, Wu K, Li S et al (2017) MFAP5 and TNNC1: potential markers for predicting occult cervical lymphatic metastasis and prognosis in early stage tongue cancer. Oncotarget 8:2525–2535