Methods

PET/CT Protocol

All PET/CT studies that were performed at our institution using either a Discovery LS (2-dimensional) or a Discovery VCT (3-dimensional) PET/CT (GE Healthcare, Milwaukee, WI). All patients who underwent PET/CT studies at our institution were scanned using a dedicated head and neck protocol as part of the established, standard clinical protocol at our institution. Patients were scanned from mid thigh to chin (body images with arms up) and then from carina to skull vertex (head and neck images obtained with arms down). The acquisition time per bed position was 5 min (2-dimensional) or 4 min 15 s (3-dimensional), respectively. The average patient blood glucose level was 97 mg/dL (range, 70-178 mg/dL). Patients were injected with an average of 575.4 ± 135.8 MBq (15.6 ± 3.7mCi) of ^{18}F-FDG, and the mean uptake time was 79 ± 19.8 min.

The PET images were reconstructed using ordered-subsets expectation maximization. The 2-dimensional implementation on the Discovery LS used 2 iterations, 28 subsets, a 5.5-mm post reconstruction Gaussian filter, and 3.9-mm pixels. The 3-dimensional implementation on the Discovery VCT (RX) used 2 iterations, 21 subsets, a 3.0-mm post reconstruction Gaussian filter, and 4.7-mm pixels. All PET data was reconstructed with and without CT-based attenuation correction. Helical CT images were obtained with a matrix of 512 × 512. Beam collimation was 10 mm with a pitch of 0.984. Tube voltage was 120 kVp, and current varied between 20 and 200 mA with auto

modulation and a noise index of 8.0. Slice thickness was 3.75 mm, and field of view was 50 cm.

Accuracy Measures

We established the overall accuracy of the scoring system by comparison with histopathologic and/or clinical confirmation of overall presence or absence of disease at 6 months of clinical follow-up from the date of the PET/CT study, as the reference standard. Imaging and electronic medical records were reviewed to determine the clinical disease status at this time point.

Outcome Measures

The primary patient outcome measures evaluated in this study were overall survival (OS) and progression free survival (PFS). A public registry of death [17] and the institution electronic medical records, until August 2013, were used to evaluate OS in this study. OS time was defined as the time from the post-treatment PET/CT study to the time of death. The data was censored for the last patient encounter in our institution. PFS time was defined as the time from the post-treatment PET/CT study to the time of clinical or pathologic recurrence, progression, or death. The data was censored for the last patient oncology visit in our institution.

Added value to clinical assessment

We also investigated whether the scoring system added value to clinical assessment. For each of the PET/CT studies, we determined whether the PET/CT study

was requested because of clinical suspicion of residual tumor at the time of the study or whether the study was ordered as part of a post therapy assessment without prior clinical suspicion or clinical assessment. This was ascertained from the indication for the study as stated in the PET/CT as written on the reports and from review of the electronic medical records, before the date of the PET/CT study.

Statistical Analysis

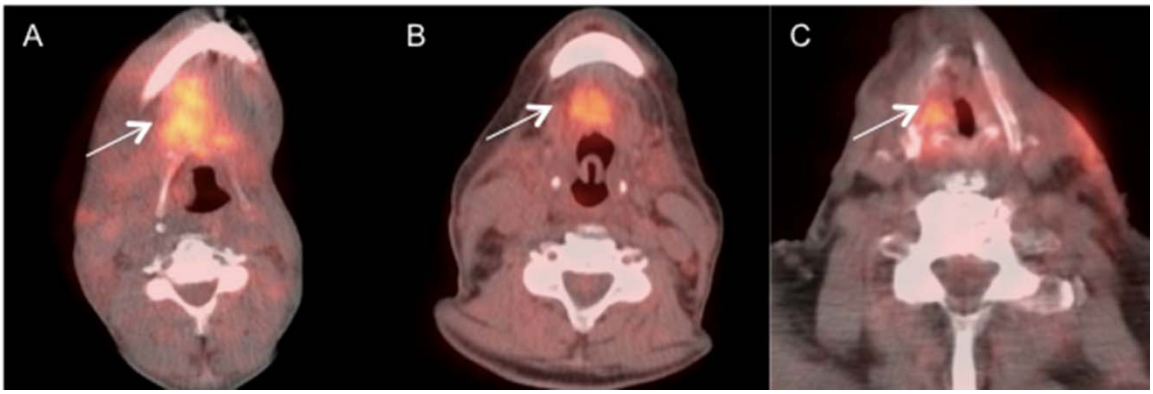
We present central tendencies as mean ± standard deviation (SD), as median (range) for data showing skewed distribution, or frequency and percentage for categorical variables. Independent samples t test or Mann –Whitney U test was used for comparing groups. Cohen's Kappa coefficient was calculated to measure inter reader reliability. The kappa values range between 0 and 1.0, with values closer to 1.00 representing better agreement between the readers. Kappa value is interpreted as follows: 0.0-0.20 indicates slight agreement, 0.21-0.40 indicates fair agreement, 0.41-0.60, indicates moderate agreement, 0.61-0.8 indicates substantial agreement, and >0.8 indicates almost perfect agreement[18]. Overall survival and progression free survival were presented as Kaplan Meier plots and compared using log-rank test. Cox multivariate analyses were performed to adjust for known important clinical factors such as age, gender, primary site, stage, HPV status, and treatment modality. We used the Prism 5 (GraphPad software Inc, San Diego) and SPSS 20 (SPSS Inc, Chicago) statistical packages for all analyses, and all statistical tests were two-sided with a significance level of 0.05.

Results

Cox multivariate regression models and survival outcomes

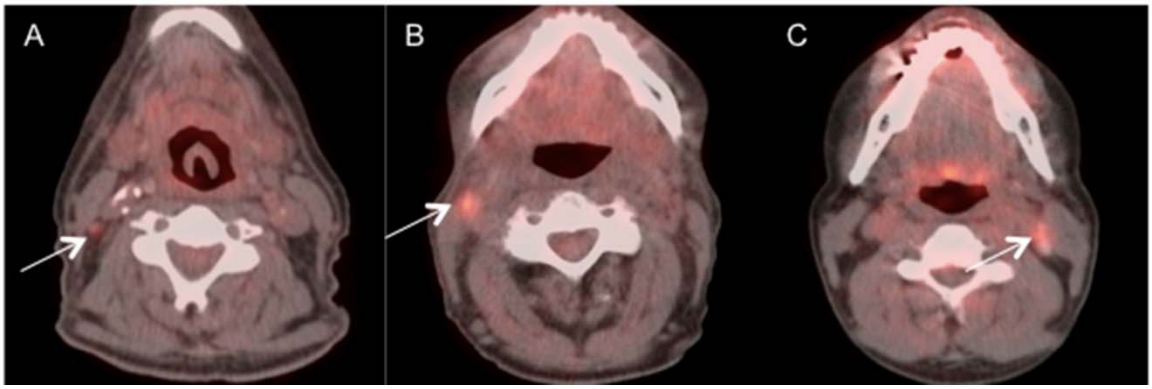
Hazard ratios (HRs) were estimated for age, male gender, HPV status, alcohol, smoking, stage (stage IV versus other), primary site (oropharynx versus other), treatment modality (Radiation therapy versus other) and PET/CT interpretation (negative for tumor vs. positive for tumor) using Cox hazard model for overall survival and progression free survival. In the univariate analysis for overall survival, age (HR, 1.04, 95%CI – 1.00-1.07; $p<0.046$), smoking (HR, 0.29, 95%CI – 0.12-0.70; $p<0.006$), primary site (oropharyngeal versus other) (HR, 4.52, 95%CI – 2.26-9.02; $p<0.0001$), HPV status (HR, 11.49; 95%CI – 3.48-37.95; $p<0.0001$), radiation therapy (HR, 2.40; 95%CI – 1.10-5.24; $p<0.029$) and PET result (HR, 0.16; 95%CI – 0.08-0.30; $p<0.0001$) were significantly associated with overall survival. In the multivariate analyses for overall survival, PET/CT result ($p<0.017$; HR, 0.14; 95%CI, 0.03-0.71) and HPV status ($p<0.009$; HR, 9.53; 95%CI, 1.18-51.93) were the only variables that were significantly associated with the time to overall survival adjusted for all other covariates in the model.

In the univariate analyses for progression free survival, smoking (HR, 0.42; 95% CI, 0.23-0.78; $p < 0.005$), primary site (HR, 2.01; 95%CI, 1.22-3.30; $p <0.006$), HPV status (HR, 4.09; 95%CI, 1.99-8.43; $p<0.0001$) and PET result (HR, 0.24, 95%CI, 0.14-0.39; $p<0.0001$) were significantly associated with progression free survival. In the multivariate analyses for progression free survival, PET/CT result ($p<0.020$; HR, 0.36; 95%CI, 0.15-0.85) and HPV status ($p<0.0001$; HR, 6.62; 95%CI, 2.41-18.18) were the only variables that were significantly associated with progression free survival.



Supplemental figure 1: Hopkins Head and Neck therapy assessment criteria:

Primary tumor – Score 3. Axial fused PET/CTs. **(A)** demonstrates diffuse FDG activity in the supraglottis, most likely post-inflammatory in nature of a 53 year old gentleman with SCC of the larynx, post chemoradiation, who underwent a post-therapy FDG PET/CT study, 14 weeks after completion of treatment. He has been on regular follow-up since the study with no evidence of active disease. **(B)** demonstrates diffuse FDG activity at the level of the hypopharynx, most likely post-inflammatory in nature of a 82 year old gentleman with SCC of the larynx, post chemoradiation, who underwent a post-therapy FDG PET/CT study, 14 weeks after completion of treatment. He has been on regular follow-up since the study with no evidence of active disease. **(C)** demonstrates diffuse FDG activity at the laryngeal soft-tissue of a 70 year old gentleman with SCC of the supraglottic larynx, post chemoradiation, who underwent a post-therapy FDG PET/CT study, 11 weeks after completion of treatment. Biopsy of the region demonstrated atypical squamous cells, consistent with residual disease.



Supplemental Figure 2: **Hopkins Head and Neck therapy assessment criteria: Neck node – Score 3.** Axial fused PET/CTs. **(A)** demonstrates mild FDG activity in the right neck, most likely post-inflammatory in nature of a 53 year old gentleman with SCC of the right tonsil, post chemoradiation, who underwent a post-therapy FDG PET/CT study, 13 weeks after completion of treatment. He has been on regular follow-up since the study with no evidence of active disease. **(B)** demonstrates diffuse FDG activity in a right level 2 neck node of a 63 year old gentleman with SCC of the right palatine tonsil, post chemoradiation, who underwent a post-therapy FDG PET/CT study, 11 weeks after completion of treatment. He underwent neck dissection and the pathology showed no evidence of disease. He has been on regular follow-up since the study with no evidence of active disease. **(C)** demonstrates diffuse FDG activity a left level 2 neck node of a 44 year old gentleman with SCC of the left tongue base, post chemoradiation, who underwent a post-therapy FDG PET/CT study, 10 weeks after completion of treatment. He underwent biopsy of the node, which was positive for active disease.