**SUPPLEMENTAL MATERIALS**

**Contents**

**Tables**

* **eTable 1. Post-training transition probabilities stratified by stage and HPV-status**
* **eTable 2. Performance of Reimbursement-Based Schedule (RBS) vs Optimized Regimens**
* **eTable 3. Modified Stage IVA HPV-positive Cohort Performance**
* **eTable 4: Python Packages Used in Model Development**

**Figures**

* **eFigure 1.** **Flowchart of model training.**
* **eFigure 2. Recurrence Model Training Algorithm**
* **eFigure 3A. Test characteristics of PET-CT scan for recurrent head and neck disease.**
* **eFigure 3B. Test characteristics of CT scan for recurrent head and neck disease**
* **eFigure 4. Comparison between training cohort (NCDB) and external validation cohort (ICON-S).**
* **eFigure 5. Cohort-specific comparison of overall survival**
* **eFigure 6. Disease-free survival stratified by stage and HPV-status and disease state (locoregional and distant metastasis)**
* **eFigure 7. Model based recommendations adapt to minimize latency based on overall number of scans.**
* **eFigure 8. Modified Stage IVA HPV-positive cohort compared to ICON-S counterpart**

**Tables**

|  |
| --- |
| **eTable 1. Post-training transition probabilities stratified by stage and HPV-status** |
| Stage | HPV-status | Year-1 | Year-2 | Year-3+ |
| P(LR) | P(mets) | P(LR) | P(mets) | P(LR) | P(mets) |
| III | + | .093% | .13% | .093% | .13% | .093% | .13% |
| - | 1.3% | .38% | 1.3% | .38% | .17% | .013% |
| IVA | + | .29% | .41% | .29% | .41% | .012% | .17% |
| - | 1.7% | .51% | 1.7% | .51% | .73% | .16% |
| IVB | + | .88% | 1.2% | .88% | 1.2% | .18% | .24% |
| - | 4.1% | 1.2% | 2.9% | .88% | .84% | .062% |
| P(LR): per month probability of locoregional recurrenceP(mets): per month probability of metastatic disease |

|  |
| --- |
| **eTable 2. Performance of Reimbursement-Based Schedule (RBS) vs Optimized Regimens** |
| cohort | RBS | optimized |
|  | Sensitivity | Mean Latency | False Positives | Sensitivity | Mean Latency | False Positives |
| III HPV+ | .58\* | 9.2‡  | 4223 | .52\* | 8.3‡  | 4277 |
| IVA HPV+ | .59\* | 10.9‡  | 3673 | .57\* | 9.4‡  | 3714 |
| IVB HPV+ | .61\* | 11.4‡  | 2623 | .63\* | 9.2‡  | 2579 |
| III HPV- | .61\* | 11.4‡  | 3012 | .67\* | 8.9‡  | 2999 |
| IVA HPV- | .61 | 10.6‡  | 2398 | .61 | 8.8‡  | 2457 |
| IVB HPV- | .66\* | 10.5‡  | 1343 | .69\* | 8.7‡  | 1394 |
| \* Denotes sensitivities that are significantly different (z score for population proportions, alpha=0.0083)‡ Denotes latencies that are significantly different (unpaired t-test, alpha=0.0083).Number of false positives was not significantly different across all cohorts (z score for population proportions, alpha=0.0083) |

|  |
| --- |
| **eTable 3. Modified Stage IVA HPV-positive Cohort Performance** |
| Regimen | Months | Sensitivity | Latency (months) | Total false positives per 10,000 patients |
| PET | 3 | .12 | 18.2 | 1111 |
| +1 CT | 3,18 | .31 | 14.6 | 1936 |
| +2 CT | 3,12,22 | .44 | 12.1 | 2770 |
| +3 CT | 3,8,14,23 | .50 | 10.3 | 3443 |
| +4 CT | 3,8,14,20,26 | .58 | 9.4 | 4035 |
| +5 CT | 3,7,12,17,22,27 | .62 | 8.2 | 4519 |
| +6 CT | 3,7,11,15,19,23,31 | .67 | 7.7\* | 4935 |
| Standard | 3,6,9,12,18,24,36 | .68 | 8.5\* | 5094 |
| Standard refers to a PET at month 3, CT neck/chest at 6,9,12,18,24,36. Latency for radiologically discovered disease is defined: latency=month of radiologic disease discovery – month of recurrence onset; Latency for radiologically missed disease is defined: latency= 36 – month of recurrence onset\* Denotes when there is a significant difference between latency of PET+6CT and Standard regimens (unpaired t-test, alpha=0.0083). There were no significant differences in sensitivity or false positives between these regimens across all cohorts (z score for population proportions, alpha=0.0083).  |

|  |
| --- |
| **eTable 4. Python Packages Used in Model Development** |
| Package Name | Purpose |
| *numpy* | Array data structure implementation |
| *pandas* | Data input and output; Data manipulation |
| *scikit-learn* | Model training and testing |
| *seaborn* | Data visualization |

**Figures**

****

**eFigure 1.** **Flowchart of model training.**

NCDB=National Cancer Database; ICON-S refers to the 2016 International Collaboration on Oropharyngeal cancer Network for Staging manuscript by O’Sullivan et al27.



**eFigure 2. Recurrence Model Training Algorithm**

**eFigure 3A. Test characteristics of PET-CT scan for recurrent head and neck disease.** LR=locoregional recurrence, red line=pooled mean value. Test characteristics given in the form of pooled mean[95% confidence interval]



**eFigure 3B. Test characteristics of CT scan for recurrent head and neck disease.** CT Neck test characteristics were used in the simulation for detection of locoregional recurrence. CT Chest test characteristics were used in the simulation for detection of metastatic recurrence. Red line=pooled mean value. Test characteristics given in the form of pooled mean[95% confidence interval]



**eFigure 4. Comparison between training cohort (NCDB) and external validation cohort (ICON-S).**

All p-values refer to log-rank test between cohorts of interest.



**eFigure 5. Cohort-specific comparison of overall survival**

(A) Comparison of the modified base model (trained model) and the training NCDB cohort. (B) External validation of trained model using ICON-S cohorts. Comparison using log-rank test.



**eFigure 6. Disease-free survival stratified by stage and HPV-status and disease state (locoregional and distant metastasis)**

****

**eFigure 7. Model based recommendations adapt to minimize latency based on overall number of scans.**

Regimen latency was normalized using the log(1+z-score- (minimum z-score of all latencies)). Darker colors correspond to greater latency. Each block represents one month. Each yellow bar represents the timing of a single scan.

PET=regimen with only a single PET-scan at month 3; +1CT, +2CT etc. indicates how many additional scans were allowed on top of the PET at month 3

****

**eFigure 8. Modified Stage IVA HPV-positive cohort compared to ICON-S counterpart**