

# Prognostic relevance of human papillomavirus infection in anal squamous cell carcinoma: analysis of the national cancer data base

Jaymin Jhaveri<sup>1</sup>, Lael Rayfield<sup>2</sup>, Yuan Liu<sup>2</sup>, Mudit Chowdhary<sup>3</sup>, Richard J. Cassidy<sup>1</sup>, Nicholas A. Madden<sup>1</sup>, Daniel G. Tanenbaum<sup>1</sup>, Theresa W. Gillespie<sup>4</sup>, Pretesh R. Patel<sup>1</sup>, Kirtesh R. Patel<sup>1,5</sup>, Jerome C. Landry<sup>1</sup>

<sup>1</sup>Department of Radiation Oncology and Winship Cancer Institute, <sup>2</sup>Biostatistics and Bioinformatics Shared Resource, Winship Cancer Institute, Emory University, Atlanta, GA, USA; <sup>3</sup>Department of Radiation Oncology, Rush University, Chicago, IL, USA; <sup>4</sup>Department of Surgery, Emory University, Atlanta, GA, USA; <sup>5</sup>Department of Therapeutic Radiology, Yale University, New Haven, CT, USA

*Contributions:* (I) Conception and design: J Jhaveri; (II) Administrative support: J Jhaveri; (III) Provision of study materials or patients: TW Gillespie, Y Liu; (IV) Collection and assembly of data: TW Gillespie, Y Liu; (V) Data analysis and interpretation: J Jhaveri, L Rayfield, Y Liu, JC Landry, PR Patel; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

*Correspondence to:* Jaymin Jhaveri, MD, MS. Department of Radiation Oncology, Emory University, 1365 Clifton Rd NE, Room AT225, Atlanta, GA 30322, USA. Email: jaymin.jhaveri@emory.edu.

**Background:** To examine the prognostic relevance of human papillomavirus (HPV) infection for anal squamous cell carcinoma (ASCC) patients treated with chemoradiation (CRT) in the National Cancer Data Base (NCDB).

**Methods:** The 2014 NCDB was queried for non-metastatic, histologically confirmed, ASCC patients diagnosed between 2004 and 2013. Patients were required to have HPV status documented in order to be eligible. Patients were then stratified into two groups: HPV+ and HPV-. Univariate analysis (UVA) was performed using the  $\chi^2$  test for categorical covariates and ANOVA for numerical covariates. Multivariable analysis (MVA) was performed using Cox proportional hazard model for overall survival (OS). Hazard ratios (HRs) and 95% confidence intervals (CIs) were generated for each covariate. To minimize selection bias, propensity score (PS) weighting was implemented to balance OS related variables between the groups including: age, education level, stage, diagnosis year, insurance type, and agent of chemotherapy.

**Results:** A total of 1,063 patients were eligible. Patients were stratified into HPV+ (n=498, 46.8%) and HPV- (n=565, 53.2%). After PS weighting, MVA for OS showed that for men, HPV infection was associated with better OS (HR: 0.60, 95% CI: 0.38–0.96; P=0.034). However, for women, HPV infection did not significantly influence survival (HR: 1.47, 95% CI: 0.96–2.25; P=0.074).

**Conclusions:** To our knowledge, this is the largest patient series evaluating the impact of HPV infection on OS in patients with anal cancer. We found that HPV infection is associated with a statistically significant better survival for men with ASCC. In contrast, for women, HPV infection did not significantly influence survival.

**Keywords:** Human papillomavirus (HPV); anal cancer; squamous cell carcinoma; overall survival (OS); National Cancer Database (NCDB); gender

Submitted Jul 17, 2017. Accepted for publication Sep 14, 2017.

doi: 10.21037/jgo.2017.10.05

View this article at: <http://dx.doi.org/10.21037/jgo.2017.10.05>

## Introduction

Carcinoma of the anal canal is a relatively uncommon malignancy accounting for only 1.9% of all malignant tumors of the gastrointestinal (GI) system in patients in the United States (1). Although rare, the incidence of this cancer appears to be increasing, particularly in women (2). Despite the rarity of this disease, anal cancer is distinct from all other GI malignancies due to the exceptionally high response rates to chemotherapy and radiation therapy (RT). For instance, rectal cancer, which is anatomically a few centimeters proximal to the anal canal, is reported to have pathologic complete response rates of 8–35% (3-6). Whereas, anal cancer is noted to have clinical response rates of 60–90% (7-9). While histology is known to play an important role in response to treatment, human papillomavirus (HPV) infection has also been previously implicated in its success to treatment.

HPV is a non-enveloped, double stranded, deoxyribonucleic acid (DNA) virus with predilection for epithelial and mucosal cells. Although there are over 200 known HPV genotypes, HPV16 and HPV18 have been established as oncogenic genotypes. Anal squamous cell carcinoma (ASCC) has been associated with sexual transmission of HPV infection (10). The prevalence of HPV DNA in ASCC ranges from 75–100%, depending on the test method, of which HPV16 is the predominant genotype (11-13). The currently adopted HPV-related oncogenic pathway, elucidated by studying cervical and head & neck squamous cell carcinomas (HNSCC), involves the viral proteins E6 and E7 which inhibit human tumor suppressor proteins p53 and Rb, respectively. Pre-clinical data of HPV infected tumor cells have demonstrated inherent sensitivity to chemotherapy and RT (14,15). In parallel, clinical series involving both HNSCC and cervical cancer HPV infected tumors have shown superior treatment response to chemoradiation (CRT) when compared to similar HPV negative tumors (16-19). Over the last decade, efforts have been made towards understanding the prognostic and predictive properties of HPV infection in HNSCC which have culminated in significant changes in the staging of HNSCC as well as on-going treatment de-escalation clinical trials.

There have been prior single institutional, retrospective studies investigating the relevance of HPV infection in ASCC, but most of these reports included a limited number of patients ranging from 47 to 153 patients (20-23). To that end, the present study was designed to answer that question

with an amplified statistical power in a multi-institutional setting by utilizing the large, prospectively acquired, National Cancer Data Base (NCDB) which captures approximately 70% of all malignancies in the United States. We sought to evaluate the impact of HPV infection on OS for patients with ASCC treated with definitive concurrent CRT.

## Methods

### *Patient selection*

The NCDB is jointly maintained by the American College of Surgeons and the American Cancer Society and includes more than 1,500 Commission on Cancer (CoC)-approved hospitals in the United States. The 2014 NCDB Participant User File (PUF) for anal cancer was used to select patients for this study. This file includes patient demographics, socioeconomic factors, disease characteristics, treatment details and survival outcomes.

The database was queried for patients diagnosed with anal cancer from 2004 to 2013. A histologic confirmation of malignancy was required in order to be eligible for the study. Patients were only included if they had a confirmed HPV infection status (either negative or positive). Patients with in-situ disease, non-squamous histology, metastatic disease at diagnosis, and disease involving the cloacogenic zone or perianal skin were excluded. We further excluded patients that did not receive definitive concurrent CRT, inappropriate RT doses or volumes, and cases with missing outcomes. Patients treated with linear accelerator radiosurgery, Gamma knife radiosurgery, brachytherapy, radium, and radioisotopes were excluded. The remaining patients were stratified into two groups: HPV+ and HPV- patients.

### *Patient demographics*

The following demographic variables were included in the analysis: age at diagnosis, gender, race, insurance status, education, geographic location, median income quartile, and treatment facility type. Patient's primary insurance carrier at the time of initial diagnosis was also evaluated. Patients were classified as having no insurance, private insurance, Medicaid, Medicare, other government insurance, or unknown insurance status. Geographic location was determined by the zip code of the patient recorded at the time of diagnosis and then it was classified and compared as metropolitan, urban, or rural location. Treatment facility was categorized as academic/research

center, which includes National Cancer Institute (NCI) designated comprehensive cancer center, or non-academic which includes community cancer program (more than 100 but  $\leq 500$  of new annual cancer cases) and comprehensive community cancer program (more than 500 new annual cancer cases). Charlson-Deyo Score was used as a surrogate for patient co-morbidities (24).

### **Disease characteristics**

The following tumor related variables were evaluated: year of diagnosis, grade (well differentiated, moderately differentiated, poorly differentiated, or undifferentiated), presence or absence of lymphovascular space invasion, clinical tumor and nodal stage based on American Joint Committee on Cancer (AJCC) staging. Patients were excluded if they had in-situ disease, T1N0 disease, or lacked histologic confirmation of malignancy. Patients with Tx clinical stage were included if they had histologic confirmation of malignancy from a nodal biopsy.

### **Treatment details**

All patients were required to have received concurrent CRT (defined as starting within 2 weeks of each modality) in order to be eligible. Radiation dose, radiation modality [intensity modulated radiation therapy (IMRT) *vs.* other], and multi-agent *vs.* single-agent chemotherapy were all evaluated. Patients were excluded if they received inappropriate RT dose ( $<46$  or  $>70$  Gy), inappropriate RT volume (outside the pelvis), or an incomplete course of RT.

### **Outcome**

The primary outcome of this study was OS—defined as time from diagnosis to time of death or last follow-up.

### **Statistical analysis**

All statistical analysis was performed using SAS 9.4 (Cary, NC, USA). Univariate associations between each variable and the study cohorts were found using the  $\chi^2$  test for categorical covariates and ANOVA for numerical covariates. The univariate association between each covariate of interest and the outcome [overall survival (OS) in months since date of diagnosis] was assessed using Cox proportional hazard model and log-rank test. A multivariable Cox proportional

hazard model for OS was fit using the backward selection method and a removal criterion of 0.20. The hazard ratios (HR) with associated 95% confidence intervals (CI) were estimated for each of the covariates and their influence on each patient group. Kaplan-Meier plots were generated to compare the survival curves of each patient group.

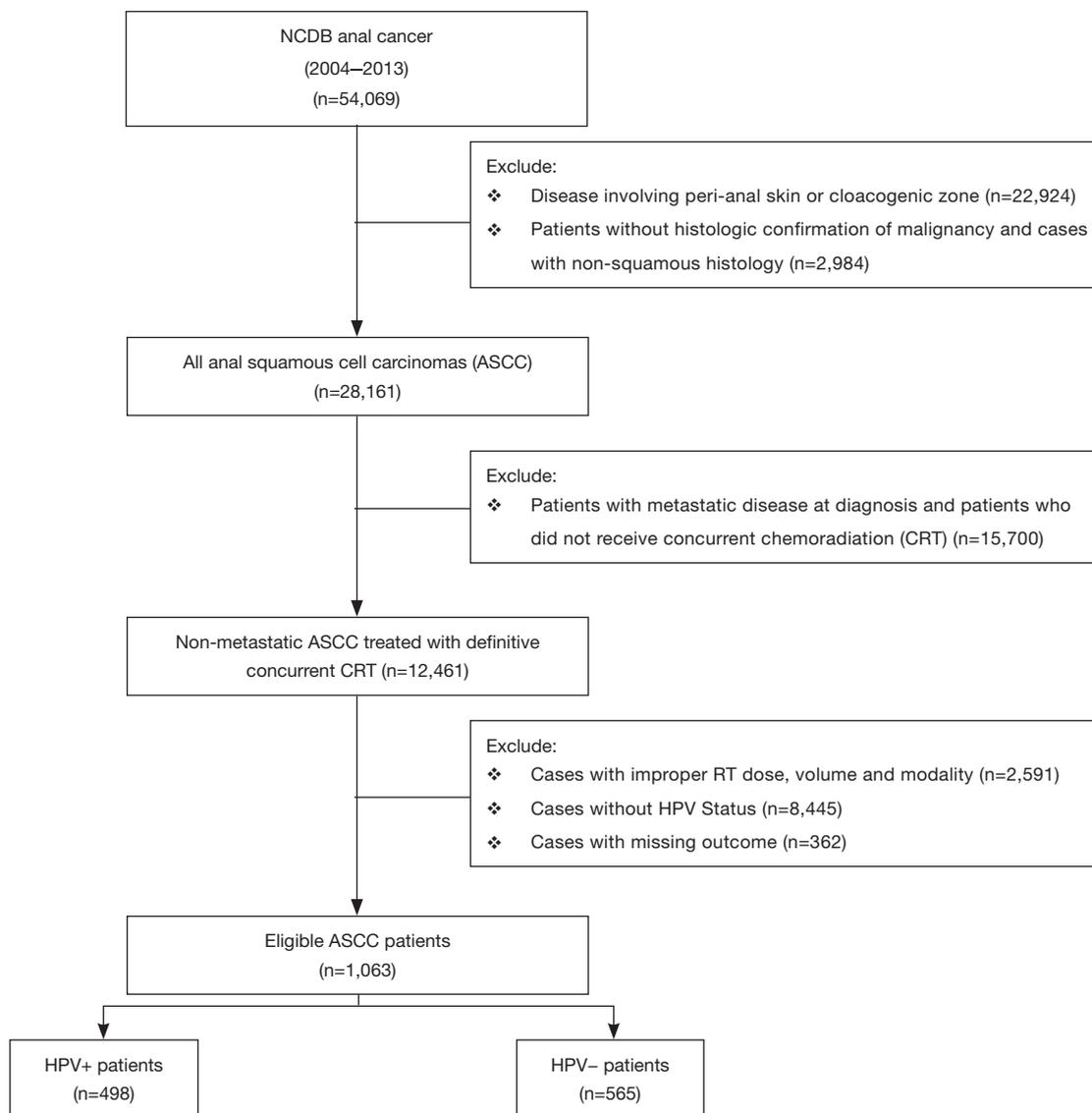
Propensity Score (PS) weighting was implemented in order to reduce the inherent imbalances between the groups (25). Due to the strong interaction between gender and OS, patients were divided into four groups for the purposes of PS weighting: HPV+ male, HPV- male, HPV+ female, and HPV- female. A multinomial logistic regression model was created to estimate the propensity of each group (26). Inverse probability of treatment weights (IPTW) estimates were calculated from the PSs and were further stabilized by multiplying them by the marginal probability of receiving the treatment observed (27). For all analyses, the weights were normalized to add up to the original sample size. The effectiveness of the weighting was evaluated by calculating the standardized differences of the covariates between treatment groups (28,29). After PS weighting was applied, the effect of HPV infection was calculated using the IPTW method with a Cox model. Weighted survival curves were generated comparing the effect of HPV infection using the Breslow method (30).

## **Results**

After applying the inclusion and exclusion criteria, there were a total of 1,063 eligible patients. Patients were stratified into two groups: HPV+ (n=498, 46.8%) and HPV- (n=565, 53.2%) as depicted in the CONSORT diagram (Figure 1). Table 1 shows detailed patient demographic, disease characteristics, and treatment information within the two groups. The median follow-up time for all patients was 32.4 months.

### **Patient characteristics**

Out of the total 1,063 eligible patients, 718 were female (67.5%) and 345 were male (32.5%). The median age at diagnosis for all patients was 57 years. There were a total of 547 patients with stage II disease (51.5%) and 516 patients with stage III disease (48.5%). A significant majority of patients received multi-agent chemotherapy (n=938, 88.2%). The mean and median RT dose were 54.03 and 54 Gy, respectively.



**Figure 1** CONSORT diagram.

### **Univariate analysis (UVA)**

UVA between the two groups showed that HPV+ patients were more likely to be male (37.4% *vs.* 26.9%,  $P < 0.001$ ), younger (median age, 55 *vs.* 58 years,  $P < 0.001$ ), have advanced clinical stage at diagnosis (stage III disease 51.9% *vs.* 44.8%,  $P = 0.021$ ), treated with single agent chemotherapy (9.2% *vs.* 6.02%,  $P = 0.010$ ), and have a Charlson-Deyo score of greater than 1 (22.7% *vs.* 17.7%,  $P = 0.044$ ). There was no statistically significant difference in race, geographic location, treatment facility type, health insurance type, median income quartile, year of diagnosis,

tumor grade, AJCC clinical tumor and nodal stage, RT dose, and RT modality between the two groups.

### **Multivariable analysis (MVA)**

Unadjusted MVA for OS showed that HPV infection was not statistically significant for all patients (HR: 0.99, 95% CI: 0.71–1.37;  $P = 0.936$ ). Male gender (HR: 1.71, 95% CI: 1.22–2.40;  $P = 0.002$ ), clinical stage III (HR: 1.97, 95% CI: 1.42–2.75;  $P < 0.001$ ), and Charlson-Deyo Score  $\geq 1$  (HR: 1.82, 95% CI: 1.28–2.58;  $P < 0.001$ ) were found to be statistically significant for OS. After stratification by

**Table 1** UVA of patient population stratified by HPV- and HPV+

Covariate	HPV		Parametric P value
	Negative (n=498)	Positive (n=565)	
Gender, n (Col %)			<0.001
Male	134 (26.91)	211 (37.35)	
Female	364 (73.09)	354 (62.65)	
Race, n (Col %)			0.101
White	441 (88.55)	481 (85.13)	
Other	57 (11.45)	84 (14.87)	
Grade, n (Col %)			1.000
Well/moderately differentiated	237 (47.59)	269 (47.61)	
Poorly differentiated/undifferentiated	144 (28.92)	163 (28.85)	
Cell type not determined	117 (23.49)	133 (23.54)	
AJCC clinical stage group, n (Col %)			0.021
Stage II	275 (55.22)	272 (48.14)	
Stage III	223 (44.78)	293 (51.86)	
Facility type, n (Col %)			0.835
Community network cancer program	103 (21.06)	118 (22.48)	
Comprehensive community cancer program	214 (43.76)	222 (42.29)	
Academic/research program	172 (35.17)	185 (35.24)	
Insurance type, n (Col %)			0.353
Not insured/unknown	30 (6.02)	46 (8.14)	
Private	254 (51.0)	299 (52.92)	
Medicaid	52 (10.44)	59 (10.44)	
Medicare/other government	162 (32.53)	161 (28.5)	
Median income quartiles, n (Col %)			0.881
< \$30,000	61 (12.73)	79 (14.31)	
\$30,000–\$35,999	86 (17.95)	94 (17.03)	
\$36,000–\$45,999	139 (29.02)	162 (29.35)	
\$46,000 +	193 (40.29)	217 (39.31)	
Percent no high school, degree, n (Col %)			0.591
≥29%	88 (18.37)	95 (17.21)	
20–28.9%	111 (23.17)	112 (20.29)	
14–19.9%	105 (21.92)	132 (23.91)	
<14%	175 (36.53)	213 (38.59)	

**Table 1** (continued)

Table 1 (continued)

Covariate	HPV		Parametric P value
	Negative (n=498)	Positive (n=565)	
Geographic location, n (Col %)			0.826
Metropolitan	402 (83.06)	460 (83.64)	
Urban	75 (15.5)	80 (14.55)	
Rural	7 (1.45)	10 (1.82)	
AJCC clinical T, n (Col %)			0.628
Stage 0	19 (3.83)	27 (4.81)	
Stage 1	300 (60.48)	314 (55.97)	
Stage 2	125 (25.2)	152 (27.09)	
Stage 3	44 (8.87)	56 (9.98)	
Stage X	8 (1.61)	12 (2.14)	
AJCC clinical N, n (Col %)			0.193
Stage 0	286 (57.55)	286 (51.07)	
Stage 1	55 (11.07)	77 (13.75)	
Stage 2	95 (19.11)	108 (19.29)	
Stage 3	57 (11.47)	82 (14.64)	
Stage X	4 (0.8)	7 (1.25)	
Agent of chemotherapy, n (Col %)			0.010
Agent not documented	28 (5.62)	15 (2.65)	
Single-agent	30 (6.02)	52 (9.2)	
Multi-agent	440 (88.35)	498 (88.14)	
Diagnosis year, n (Col %)			0.409
2008–2011	186 (37.35)	200 (35.4)	
2012	149 (29.92)	158 (27.96)	
2013	163 (32.73)	207 (36.64)	
Charlson-Deyo Score, n (Col %)			0.044
0	410 (82.33)	437 (77.35)	
1+	88 (17.67)	128 (22.65)	
Radiation treatment modality, n (Col %)			0.823
Other	174 (34.94)	206 (36.46)	
IMRT	302 (60.64)	337 (59.65)	
Conformal/3D therapy	22 (4.42)	22 (3.89)	

Table 1 (continued)

Table 1 (continued)

Covariate	HPV		Parametric P value
	Negative (n=498)	Positive (n=565)	
Age at diagnosis			<0.001
N	498	565	
Mean	59.42	55.49	
Median	58	55	
Min	24	23	
Max	89	90	
Standard deviation	10.74	11.13	
Radiation dose (cGy)			0.901
N	498	565	
Mean	5,401.63	5,405.01	
Median	5,400	5,400	
Min	4,400	4,500	
Max	6,840	6,840	
Standard deviation	464.19	422.84	

UVA, univariate analysis; IMRT, intensity modulated radiation therapy; HPV, human papillomavirus; AJCC, American Joint Committee on Cancer.

gender, MVA showed that presence of HPV infection was a statistically significant variable for men (HR: 0.60, 95% CI: 0.39–0.94;  $P=0.025$ ). Furthermore, there was a statistically significant difference in OS between men and women (type III  $P=0.006$ ).

### Unadjusted OS

The unadjusted Kaplan-Meier analysis for OS (Figure 2) showed a 5-year survival of 72.7% vs. 75.3% for HPV+ and HPV- groups, respectively ( $P=0.5978$ ). With further stratification based on gender, unadjusted Kaplan-Meier analysis for OS showed that within male patients, HPV+ men had improved 5-year OS when compared to HPV- men (68.1% vs. 48.3%,  $P=0.0379$ ) as depicted in Figure S1. For female patients, HPV+ women had a trend towards worse 5-year OS when compared to HPV- women (78.4% vs. 86.2%,  $P=0.0670$ ) as depicted in Figure S2.

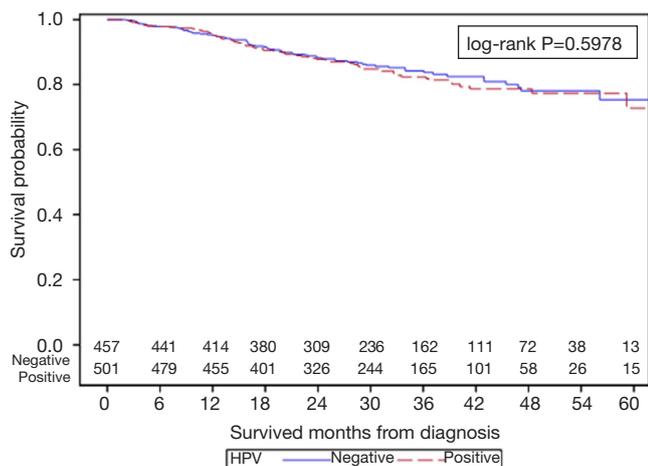
### PS analysis

Due to the strong interaction between gender and OS,

patients were divided into four groups: HPV+ male (n=180, 18.8%), HPV- male (n=120, 12.5%), HPV+ female (n=321, 33.5%), and HPV- female (n=337, 35.2%). PS weighting was applied across the four groups. Table S1 shows the balance check between the four PS weighted groups. The standardized difference for known co-variables including age, clinical stage, treatment facility type, type of insurance, year of diagnosis, Charlson-Deyo score, and chemotherapy agent were all calculated to be <0.1. For the diagnosis year 2012, the standardized differences amongst all four groups slightly exceeded 0.1, however, this was considered to be statistically and clinically insignificant.

PS weighted MVA for OS stratified by gender shows HPV+ men have a statistically significant improved OS when compared to HPV- men (HR: 0.60, 95% CI: 0.38–0.96;  $P=0.034$ ). HPV+ women demonstrated a statistical trend towards worse OS when compared to similar HPV- women (HR: 1.47, 95% CI: 0.96–2.25;  $P=0.074$ ).

PS weighted KM curves for all four groups are shown in Figure 3. This confirms that HPV+ men have improved 5-year OS of 56.4% when compared to HPV- men of 50.0% ( $P=0.0340$ , Figure S3) and HPV+ women



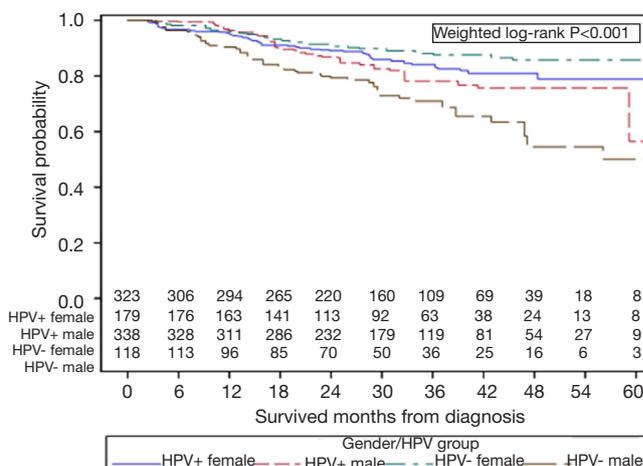
**Figure 2** Unadjusted Kaplan-Meier survival analysis of HPV+ and HPV- patients. HPV, human papillomavirus.

demonstrate a trend towards inferior 5-year OS of 78.9% vs. 85.6% when compared to HPV- women (P=0.0740, Figure S4).

**Discussion**

Our investigation shows that HPV infection is a favorable prognostic factor for men with ASCC treated with definitive CRT. However, HPV infection conferred a statistical trend towards worse OS for women with ASCC. To our knowledge, the present study with 1,063 patients is the largest reported series that evaluates the impact of HPV infection in ASCC in a multi-institutional setting. Furthermore, the finding that HPV infection has a differential prognostic effect based on gender, has never been reported.

The finding that HPV infection portends an improved prognosis in male ASCC patients treated with CRT is congruent with previously published reports. Yhim *et al.* found that 31 HPV infected ASCC patients treated with definitive CRT had superior progression free survival, time to local failure, and OS when compared to 16 HPV negative patients (20). Similarly, Mai *et al.*, using p16 as a surrogate marker for HPV infection, reported that 137 p16+ ASCC patients treated with CRT had improved relapse free survival and cancer specific survival when compared to 16 p16- patients (22). However, none of the previously reported studies separated the effect of HPV infection on gender. Initial results of RTOG 98-11 demonstrated that on MVA, male sex (P=0.02) was an independent prognosticator



**Figure 3** Adjusted Kaplan-Meier survival analysis of four cohorts stratified by gender and HPV infection-. HPV, human papillomavirus.

for worse disease free-survival (31). This difference in survival persisted in the long-term report of RTOG 98-11 where male patients had statistically worse OS (HR: 1.38, P=0.031) (8). Our analysis aligns with the finding of RTOG 98-11 where male gender is independently associated with statistically significant inferior OS (HR: 1.71, 95% CI: 1.22–2.40; P=0.002). The reason for this disparity in survival for ASCC patients still remains unknown. The gender specific analysis is particularly important because the opposing effect of HPV infection on gender leads to an initial observation of no difference in OS with overlapping KM survival curves (Figure 1). Only when the analysis is separated by gender (Figures S1-4), were we able to discover the differing effect of HPV infection on gender, which explains the initial observation of no difference in OS.

Despite the application of rigorous statistical techniques to minimize selection bias, the differential effect of HPV infection on gender still persisted on MVA. Although there was a statistical trend towards worse OS for HPV+ women, the divergent effect of HPV infection on gender is difficult to explain. In theory, HPV infection yields an increased sensitivity to CRT due to multiple factors: increased levels of excision repair gene expression (32), elevated RT induced residual DNA double strand breaks (DSBs) (14), modulation of protein kinase B activation (33), and restoration of apoptotic cell death and upregulation of tumor suppressor p53 (15). The underlying oncogenic pathway and radiosensitivity due to HPV infection in squamous cell cancers has not been shown to be different

based on gender. However, emerging data indicates that for HPV associated cancers, there is a high risk subgroup with elevated levels of E6 gene expression that is at increased risk of distant metastases and demonstrates worse cancer specific survival (34).

Our report adds to the increasing amount of literature that suggests HPV infection has a prognostic role in ASCC. However, our study is unique as we were able to validate the impact of HPV infection on OS in a large, multi-institutional database. Moreover, the gender specific impact of HPV infection has never been previously reported. Based on the results of our investigation, and other prior studies, the impact of HPV infection in ASCC should be investigated in a prospective, randomized clinical trial in order to eliminate the inherent selection bias associated with retrospective analyses. Furthermore, patients should be stratified based on gender, as the present study and RTOG 98-11 reported improved survival for women with ASCC.

Our study has a few pertinent limitations. HPV infection was categorized as being positive or negative—the method of testing (in-situ hybridization or polymerase chain reaction) was not available in the NCDB. Of the HPV positive patients, majority of them had missing genotype information and hence a subset analysis of each HPV genotype could not be performed. Moreover, p16 overexpression—a surrogate marker for HPV infection routinely used in the clinic—was unavailable for analysis. Therefore, we were unable to stratify and analyze patients based on p16 overexpression. Next, the NCDB does not specify chemotherapy agents (mitomycin versus cisplatin) and this could not be included in the analysis. Lastly, human immunodeficiency virus (HIV) infection status was also not available in the NCDB and hence could not be included in the analysis.

## Conclusions

HPV infection was associated with a survival advantage for male ASCC patients treated with definitive CRT. However, HPV infection did not influence survival for similar female ASCC patients treated with definitive CRT. The differential effect of HPV infection on gender in ASCC is a novel finding which warrants further investigation in prospective, randomized clinical trials.

## Acknowledgements

We would like to thank the American College of Surgeons

Commission on Cancer and the American Cancer Society for access to the data that enabled this analysis.

*Funding:* Research reported in this publication was supported in part by the Biostatistics and Bioinformatics Shared Resource of Winship Cancer Institute of Emory University and NIH/NCI award number P30CA138292.

## Footnote

*Conflicts of Interest:* The authors have no conflicts of interest to declare.

*Disclaimer:* The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

## References

1. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2017. *CA Cancer J Clin* 2017;67:7-30.
2. Howlader N NA, Krapcho M, Miller D, et al. SEER Cancer Statistics Review. Based on November 2015 SEER data submission. April 2016 ed. Bethesda, MD: National Cancer Institute, 2016.
3. Sauer R, Becker H, Hohenberger W, et al. Preoperative versus Postoperative Chemoradiotherapy for Rectal Cancer. *N Engl J Med* 2004;351:1731-40.
4. Belluco C, De Paoli A, Canzonieri V, et al. Long-Term Outcome of Patients with Complete Pathologic Response after Neoadjuvant Chemoradiation for cT3 Rectal Cancer: Implications for Local Excision Surgical Strategies. *Ann Surg Oncol* 2011;18:3686-93.
5. Maréchal R, Vos B, Polus M, et al. Short course chemotherapy followed by concomitant chemoradiotherapy and surgery in locally advanced rectal cancer: a randomized multicentric phase II study. *Ann Oncol* 2012;23:1525-30.
6. Park SH, Kim JC. Preoperative chemoradiation for locally advanced rectal cancer: comparison of three radiation dose and fractionation schedules. *Radiat Oncol J* 2016;34:96-105.
7. Northover J, Glynne-Jones R, Sebag-Montefiore D, et al. Chemoradiation for the treatment of epidermoid anal cancer: 13-year follow-up of the first randomised UKCCCR Anal Cancer Trial (ACT I). *Br J Cancer* 2010;102:1123-8.
8. Gunderson LL, Winter KA, Ajani JA, et al. Long-term update of US GI intergroup RTOG 98-11 phase III trial

- for anal carcinoma: survival, relapse, and colostomy failure with concurrent chemoradiation involving fluorouracil/mitomycin versus fluorouracil/cisplatin. *J Clin Oncol* 2012;30:4344-51.
9. James RD, Glynn-Jones R, Meadows HM, et al. Mitomycin or cisplatin chemoradiation with or without maintenance chemotherapy for treatment of squamous-cell carcinoma of the anus (ACT II): a randomised, phase 3, open-label, 2x2 factorial trial. *Lancet Oncol* 2013;14:516-24.
  10. zur Hausen H. Papillomaviruses and cancer: from basic studies to clinical application. *Nat Rev Cancer* 2002;2:342-50.
  11. De Vuyst H, Clifford G, Li N, et al. HPV infection in Europe. *Eur J Cancer* 2009;45:2632-9.
  12. Skamperle M, Kocjan BJ, Maver PJ, et al. Human papillomavirus (HPV) prevalence and HPV type distribution in cervical, vulvar, and anal cancers in central and eastern Europe. *Acta Dermatovenerol Alp Pannonica Adriat* 2013;22:1-5.
  13. Palefsky JM, Holly EA, Gonzales J, et al. Detection of human papillomavirus DNA in anal intraepithelial neoplasia and anal cancer. *Cancer Res* 1991;51:1014-9.
  14. Rieckmann T, Tribius S, Grob TJ, et al. HNSCC cell lines positive for HPV and p16 possess higher cellular radiosensitivity due to an impaired DSB repair capacity. *Radiother Oncol* 2013;107:242-6.
  15. Kimple RJ, Smith MA, Blitzer GC, et al. Enhanced radiation sensitivity in HPV-positive head and neck cancer. *Cancer Res* 2013;73:4791-800.
  16. Lassen P, Eriksen JG, Hamilton-Dutoit S, et al. Effect of HPV-associated p16INK4A expression on response to radiotherapy and survival in squamous cell carcinoma of the head and neck. *J Clin Oncol* 2009;27:1992-8.
  17. Ang KK, Harris J, Wheeler R, et al. Human papillomavirus and survival of patients with oropharyngeal cancer. *N Engl J Med* 2010;363:24-35.
  18. Rischin D, Young RJ, Fisher R, et al. Prognostic significance of p16INK4A and human papillomavirus in patients with oropharyngeal cancer treated on TROG 02.02 phase III trial. *J Clin Oncol* 2010;28:4142-8.
  19. Vosmik M, Laco J, Sirak I, et al. Prognostic significance of human papillomavirus (HPV) status and expression of selected markers (HER2/neu, EGFR, VEGF, CD34, p63, p53 and Ki67/MIB-1) on outcome after (chemo-) radiotherapy in patients with squamous cell carcinoma of uterine cervix. *Pathol Oncol Res* 2014;20:131-7.
  20. Yhim HY, Lee NR, Song EK, et al. The prognostic significance of tumor human papillomavirus status for patients with anal squamous cell carcinoma treated with combined chemoradiotherapy. *Int J Cancer* 2011;129:1752-60.
  21. Gilbert DC, Williams A, Allan K, et al. p16INK4A, p53, EGFR expression and KRAS mutation status in squamous cell cancers of the anus: correlation with outcomes following chemo-radiotherapy. *Radiother Oncol* 2013;109:146-51.
  22. Mai S, Welzel G, Ottstadt M, et al. Prognostic Relevance of HPV Infection and p16 Overexpression in Squamous Cell Anal Cancer. *Int J Radiat Oncol Biol Phys* 2015;93:819-27.
  23. Rodel F, Wieland U, Fraunholz I, et al. Human papillomavirus DNA load and p16INK4a expression predict for local control in patients with anal squamous cell carcinoma treated with chemoradiotherapy. *Int J Cancer* 2015;136:278-88.
  24. Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *J Clin Epidemiol* 1992;45:613-9.
  25. Austin PC. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behav Res* 2011;46:399-424.
  26. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med* 2009;28:3083-107.
  27. McCaffrey DF, Griffin BA, Almirall D, et al. A tutorial on propensity score estimation for multiple treatments using generalized boosted models. *Stat Med* 2013;32:3388-414.
  28. Imbens GW. The role of the propensity score in estimating dose-response functions. *Biometrika* 2000;87:706-10.
  29. Austin PC. The relative ability of different propensity score methods to balance measured covariates between treated and untreated subjects in observational studies. *Med Decis Making* 2009;29:661-77.
  30. Cole SR, Hernan MA. Adjusted survival curves with inverse probability weights. *Comput Methods Programs Biomed* 2004;75:45-9.
  31. Ajani JA, Winter KA, Gunderson LL, et al. Fluorouracil, mitomycin, and radiotherapy vs fluorouracil, cisplatin, and radiotherapy for carcinoma of the anal canal: a randomized controlled trial. *JAMA* 2008;299:1914-21.
  32. Hampson L, El Hady ES, Moore JV, et al. The HPV16 E6 and E7 proteins and the radiation resistance of cervical carcinoma. *FASEB J* 2001;15:1445-7.
  33. Gupta AK, Lee JH, Wilke WW, et al. Radiation response

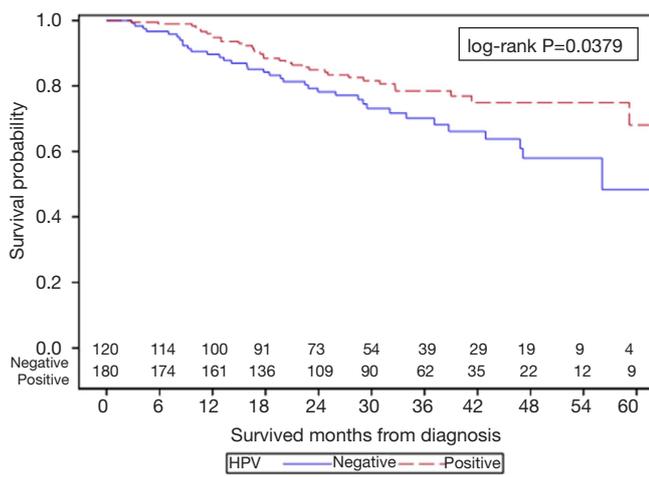
in two HPV-infected head-and-neck cancer cell lines in comparison to a non-HPV-infected cell line and relationship to signaling through AKT. *Int J Radiat Oncol Biol Phys* 2009;74:928-33.

34. Khwaja SS, Baker C, Haynes W, et al. High E6 Gene

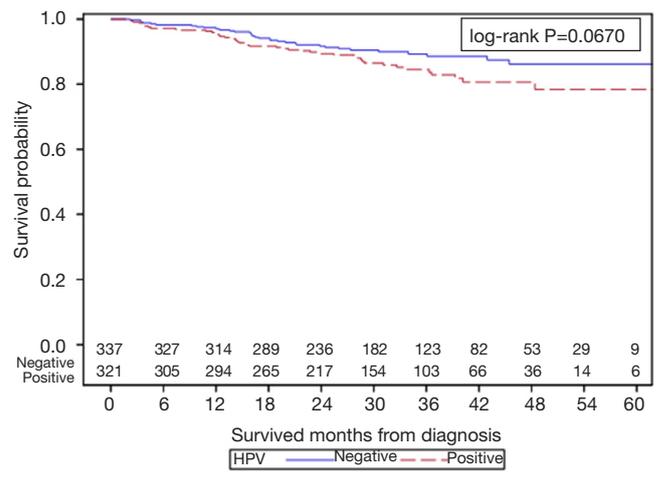
Expression Predicts for Distant Metastasis and Poor Survival in Patients With HPV-Positive Oropharyngeal Squamous Cell Carcinoma. *Int J Radiat Oncol Biol Phys* 2016;95:1132-41.

**Cite this article as:** Jhaveri J, Rayfield L, Liu Y, Chowdhary M, Cassidy RJ, Madden NA, Tanenbaum DG, Gillespie TW, Patel PR, Patel KR, Landry JC. Prognostic relevance of human papillomavirus infection in anal squamous cell carcinoma: analysis of the national cancer data base. *J Gastrointest Oncol* 2017;8(6):998-1008. doi: 10.21037/jgo.2017.10.05

Supplementary



**Figure S1** Unadjusted Kaplan-Meier survival analysis for men stratified by HPV infection. HPV, human papillomavirus.

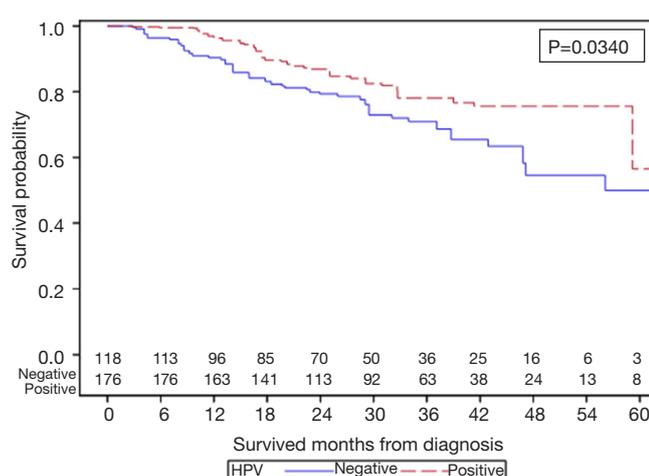


**Figure S2** Unadjusted Kaplan-Meier survival analysis for women stratified by HPV infection. HPV, human papillomavirus.

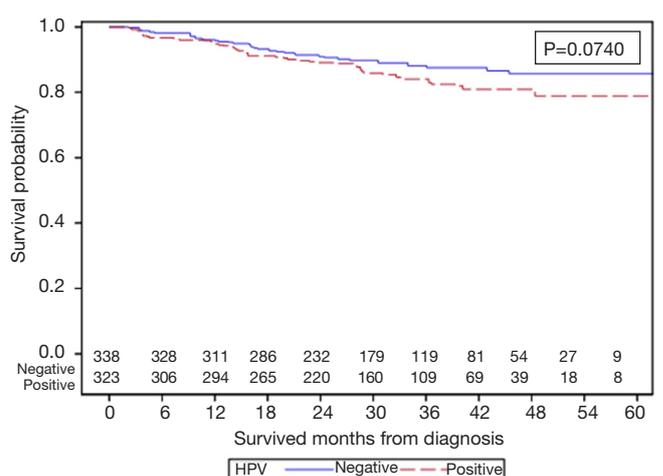
**Table S1** PS weighting of four groups: HPV+ male, HPV- male, HPV+ female, HPV- female

Covariate	Gender/HPV group				Parametric P value*	Standardized difference
	Male HPV negative (n=118)	Male HPV positive (n=179)	Female HPV negative (n=338)	Female HPV positive (n=322)		
AJCC clinical stage group, n (Col %)					0.906	
Stage II	64 (54.2)	94 (52.55)	172 (50.94)	172 (53.35)		0.065
Stage III	54 (45.8)	85 (47.45)	166 (49.06)	150 (46.65)		0.065
Facility type, n (Col %)					0.999	
Community/integrated network cancer program	27 (22.5)	41 (22.63)	72 (21.24)	69 (21.49)		0.034
Comprehensive community cancer program	49 (41.46)	77 (43.1)	147 (43.52)	139 (43.02)		0.042
Academic/research program	43 (36.04)	61 (34.26)	119 (35.24)	114 (35.49)		0.037
Primary payor, n (Col %)					0.999	
Not insured/unknown	9 (7.78)	13 (7.46)	24 (7.24)	21 (6.48)		0.050
Private	58 (49.42)	94 (52.49)	177 (52.45)	164 (51.02)		0.061
Medicaid	11 (9.35)	14 (7.9)	29 (8.57)	32 (9.91)		0.070
Medicare/other government	39 (33.45)	58 (32.15)	107 (31.74)	105 (32.59)		0.036
Urban/rural 2003, n (Col %)					0.985	
Metropolitan	103 (87.19)	152 (85.03)	286 (84.73)	269 (83.66)		0.100
Urban	13 (11.4)	25 (13.87)	47 (13.87)	48 (14.85)		0.102
Rural	2 (1.4)	2 (1.1)	5 (1.4)	5 (1.49)		0.034
Diagnosis year (quartile), n (Col %)					0.978	
2008–2011	40 (34.31)	68 (37.93)	119 (35.29)	113 (35.19)		0.075
2012	38 (32.37)	49 (27.15)	98 (29.08)	95 (29.76)		0.114
2013	39 (33.32)	63 (34.92)	120 (35.63)	112 (35.04)		0.049
Charlson-Deyo Score, n (Col %)					0.918	
0	95 (80.71)	139 (77.48)	265 (78.35)	251 (77.91)		0.080
1+	22 (19.29)	40 (22.52)	73 (21.65)	71 (22.09)		0.080
Agent of chemotherapy, n (Col %)					0.984	
Agent not documented	6 (4.72)	9 (5.05)	15 (4.54)	11 (3.53)		0.075
Single-agent	10 (8.09)	12 (6.86)	25 (7.53)	26 (8.07)		0.047
Multi-agent	103 (87.2)	158 (88.09)	297 (87.93)	285 (88.4)		0.037
Age at diagnosis (years), n (Col %)					0.970	
<50	25 (21.24)	34 (19.56)	71 (21.01)	64 (19.95)		0.042
≥50	93 (78.76)	144 (80.44)	267 (78.99)	258 (80.05)		0.042
Percent no high school degree quartiles 2000, n (Col %)					0.996	
≥29%	21 (17.66)	32 (18.03)	62 (18.21)	60 (18.58)		0.024
20–28.9%	26 (22.42)	33 (18.46)	73 (21.66)	72 (22.25)		0.098
14–19.9%	25 (21.11)	44 (24.36)	78 (23.03)	70 (21.83)		0.078
<14%	46 (38.81)	70 (39.15)	125 (37.11)	120 (37.35)		0.042

\*, the parametric P value is calculated by ANOVA for numerical covariates and Chi-Square test for categorical covariates based on inverse propensity score weighted sample. The frequency was rounded to the nearest integer, and the summation may not equal to the exact sample total. PS, propensity score; HPV, human papillomavirus.



**Figure S3** Adjusted Kaplan-Meier survival analysis for men stratified by HPV infection. HPV, human papillomavirus.



**Figure S4** Adjusted Kaplan-Meier survival analysis for women stratified by HPV infection. HPV, human papillomavirus.