Associations between human papillomavirus and history of cancer among U.S. adults in the National Health and Nutrition Examination Survey (2003–2010)

B Liu1 and E Taioli *,1

1Department of Population Health and the Feinstein Institute, Hofstra North Shore-LIJ School of Medicine, 175 Community Drive, Rm 203, Great Neck, NY 11021, USA

Background: Human papillomavirus (HPV) is an infectious agent that has been associated with human cancer. We have updated the U.S. population sero-prevalence using a large National Health and Nutrition Examination Survey (NHANES) sample of adults from 2003 to 2010, and have analysed the associations between HPV seropositivity and self-reported history of cancer.

Methods: Four cross-sectional cycles (2003–2004, 2005–2006, 2007–2008, and 2009–2010) were used, for a total of 12,759 participants who had both cancer history and HPV serum information.

Results: The sero-prevalences of HPV types 6, 11, 16, and 18 were 15.0%, 4.8%, 11.5%, and 4.1%, respectively. Females had significantly higher HPV prevalence than males (P < 0.05) for all subtypes. Positive associations between HPV 16/18 seropositivity and lifetime history of any cancer (adjusted odds ratio-OR adj = 1.68; 95% CI: 1.35, 2.01), history of any of eight selected cancers (OR adj = 2.63; 95% CI: 1.78, 3.90), lung cancer (OR adj = 5.14; 95% CI: 1.29, 20.44), and cervical cancer (OR adj = 2.55; 95% CI: 1.63, 3.98) were observed.

Conclusions: The finding of significant associations between HPV 16/18 seropositivity and lifetime history of cancer adds epidemiological evidence to the carcinogenicity potential of HPV 16 and 18 in other tissues. With increasing coverage of the HPV vaccine in the U.S., future NHANES data and sample collection may allow further detailed evaluation of the population impact of the HPV vaccination on cancer prevention.
Geijerstam et al., 1998; Markowitz et al., 2009, 2013), and the implication of HPV as an aetiologic risk factor for some major human cancers, we are also analysing the associations between HPV seropositivity and self-reported history of cancer. We hypothesised that seropositivity of HPV 16 or 18 is associated with a higher frequency of personal history of cancer compared with seronegativity.

MATERIALS AND METHODS

Study population and design. The NHANES surveys were conducted by the U.S. National Center for Health Statistics (NCHS) of the Centers for Disease Control and Prevention (CDC) to assess the health and nutritional status of adults and children in the U.S. The NHANES study protocol has been described previously in detail by CDC (Zipf et al., 2013). In brief, civilians, non-institutionalized persons in the U.S. aged 2 months or older were eligible to participate. They were selected through a stratified, multi-stage, probability-cluster design. The survey’s components were administered in homes and in mobile examination centres, and also provided national reference data for biological markers, anthropometric measures, as well as demographic and socioeconomic status data.

Four cross-sectional cycles (2003–2004, 2005–2006, 2007–2008, and 2009–2010) were used and combined using NCHS recommendations (NCHS, 2008, 2013). Of the initial 41,156 participants, those who lacked information on either only HPV serum (n = 9,223), only on history of cancer (n = 1,562), or both (n = 17,455) were excluded. Among the remaining 12,916 participants who had both cancer and HPV serum information, 45 and 112 participants from the 2007–2008 and 2009–2010 survey cycles, respectively, received HPV vaccination, and these 157 people were also excluded. The final population consisted of 12,759 adults aged 18–59 years. The NHANES protocol was developed and reviewed in compliance with the policies for protection of human research subjects developed by the U.S. Department of Health and Human Services (Zipf et al., 2013). The Institutional Review Board of the Feinstein Institute for Medical Research determined that this study is IRB exempt.

History of cancer. Lifetime cancer history was defined based on a positive answer to the question in the medical condition questionnaire: ‘Have you ever been told by a doctor or other health professional that you had cancer or a malignancy of any kind?’.

Eight specific cancers that were potentially linked to HPV were identified based on positive answers to the subsequent questions ‘What kind of cancer?’; they were cervix, oesophagus, larynx/ windpipe, lung, mouth/tongue/lip, ovary, prostate, and rectum. No information on vulvar, anal, penile, or vaginal cancers was collected. An additional dichotomised cancer variable (eight selected cancer) was created which included a positive answer to having had any of those eight cancers.

Serum HPV. All NHANES participants aged 18–59 years were eligible for serologic testing. Serum samples were analysed using the LumineX platform to simultaneously assay antibodies to HPV 6, 11, 16, and 18, and were detected by a multiplexed, competitive immunoassay by displacement of fluorescently tagged neutralising monoclonal antibody from virus-like particles (VLP)-coated microspheres (Opalka et al., 2003; Dias et al., 2005). Dichotomized serostatus (positive or negative) were defined by NHANES based on cutoff values (mMU ml⁻¹): ≥20, ≥16, ≥20, and ≥24 for types 6, 11, 16, and 18, respectively. For analytic purposes, three additional HPV variables were created: positivity to any of the four HPV types (HPV any four), HPV 6/11, and HPV 16/18, with the later two representing low- and high-risk HPV types, respectively.

Other covariates. The associations between cancer and HPV serostatus were adjusted for the following a priori possible confounders: sex, age, race, education, poverty income ratio (PIR), body mass index (BMI), smoking status, and alcohol use. Age was categorised into four groups: 18–29, 30–39, 40–49, and 50–59 years. Race included three categories: non-Hispanic white, non-Hispanic black, and other, where other consisted of mainly Mexican-American and other Hispanic, as well as other races (including multi-racial subjects). The latter category represented approximately 3–6% of the total sample in each survey cycle. Education was grouped into three categories: greater, equal, or less than high school level, where high school was defined as having 12 years of primary and secondary school. PIR in NHANES was calculated by dividing family income by the poverty level issued by the Department of Health and Human Services, which is specific to family size, the appropriate year and state. In this study, PIR was grouped into three categories: <1, 1 ≤ PIR < median, > median, where medians (~2.3–2.7) were calculated based on the PIR ≥1 for each survey cycle. BMI was grouped into two categories based on the cutoff value of 25 kg m⁻², where BMI ≥25 kg m⁻² indicates overweight and obese weight status based on commonly used health guidelines (CDC; NIH). Dichotomised smoking status was based on positive answers to the question, ‘Have you smoked at least 100 cigarettes in your entire life?’.

Statistical methods. The sero-prevalences of HPV types were calculated for the overall data, as well as stratified by sex, race, age groups, survey cycles, and cancer status. The HPV–cancer associations were tested using logistical regression, and models were fit with one cancer variable at a time. The crude model was fit with each of the seven HPV indicators as the only predictor (one at a time), to obtain a preliminary understanding of the HPV–cancer associations. The adjusted models were run with HPV 16/18 only, as 16 and 18 are high-risk viral subtypes. Models run on the whole-data set consisted of both males and females, except for cervical cancer and ovarian cancer, which were run on a female-only data set, and prostate cancer, which was run on a male-only data set.

Statistical analyses were performed using SAS software (version 9.3, Cary, NC, USA) and sampling weights were applied to account for the complex sampling NHANES design according to NHANES guidelines (NCHS, 2013).

RESULTS

Table 1 lists the weighted characteristics of the participants 18–59 years of age from NHANES 2003–2010 included in this study (n = 12,759). The weighted mean age (± s.e.) of the participants was 39.6 ± 0.16 years. Half of the participants were females and about 66% of the total population was overweight or obese (BMI ≥25 kg m⁻²). Non-Hispanic whites accounted for 68% of the total population; 11.5% were non-Hispanic blacks. Approximately 19% of the participants were from families with income below the poverty level (PIR <1), and 17% of the participants had lower than high school education. Non-smokers and non-drinkers accounted for 52% and 22% of the total population, respectively. The prevalence (± s.e.) of lifetime history of cancer, any of the eight selected cancers, and cervical cancer were 4.88 ± 0.26%, 1.44 ± 0.16%, and 0.97 ± 0.12%, respectively. The prevalence of the other cancer types was below 0.13%.

The sero-prevalences of HPV types 6, 11, 16, and 18 were 15.0%, 4.8%, 11.5%, and 4.1%, respectively, and remained relatively stable across the four survey cycles between 2003 and 2010: prevalence ranges were 12.8–17.2%, 4.4–5.4%, 11.1–11.8%, and 3.7–4.6% for HPV types 6, 11, 16, and 18, respectively. Females had significantly
higher HPV prevalence than males ($P < 0.05$) for all subtypes (Figure 1). The sero-prevalence of HPV 6/11 was 17.2% overall, and was 24.9% and 9.5% among females and males, respectively. The sero-prevalence of HPV 16/18 was 14.0% overall, and was 21.4% and 6.7% among females and males, respectively. The sero-prevalence of any of the four HPV types was 25.5% overall, and was 36.6% and 14.4% among females and males, respectively. The sero-prevalences of the HPV types also varied by age (Figure 2) and by race/ethnicity (Figure 3), with elevated sero-prevalences seen among those aged 30–39 years and among blacks.

Compared with non-cancer cases, sero-prevalences of HPV 16/18 were generally higher among participants with a lifetime history of cancer (Table 2), for history of any of the eight selected cancers, cervical, lung, and ovarian cancer. In general, the prevalence of HPV 16/18 among participants with a history of cancer was 1.6–3 times higher than the overall HPV 16/18 prevalence (14%). Uncertainties of the estimates for history of other cancers are large, likely because of the limited number of cancer cases ($n < 5$) and/or large standard errors ($>30$%)

Results from the crude logistic models showed significant positive associations between HPV 16/18 serum positivity and the following variables: lifetime history of cancer, history of any of the eight selected cancers, cervical, and lung cancers. HPV 16/18 serum positivity was not significantly associated with larynx/windpipe, mouth/tongue/lip, ovarian, and prostate cancers (data not shown). Due to the small sample size of oesophageal (two cases) and rectal cancer (one case) cases, no reliable results were found as the convergence criteria required to run the model were not met.

**Multivariable analysis.** For the study population as a whole, positive associations between HPV 16/18 seropositivity and lifetime history of any cancer (adjusted odds ratio-OR$_{adj}$ = 1.68; 95% CI: 1.35, 2.01), history of any of the eight selected cancers (OR$_{adj}$ = 2.63; 95% CI: 1.78, 3.90), and lung cancer (OR$_{adj}$ = 5.14; 95% CI: 1.29, 20.44) were observed (Table 3). Age and sex were two common significant covariates in the adjusted models for these three types of cancer. Among females, similar positive associations were found between HPV 16/18 seropositivity and history of cancer (OR$_{adj}$ = 1.57; 95% CI: 1.22, 2.01), history of any of the eight selected cancers (OR$_{adj}$ = 2.57; 95% CI: 1.64, 4.03), and lung cancer (OR$_{adj}$ = 5.04; 95% CI: 1.34, 18.92). In addition, HPV 16/18 was found to be a significant predictor for cervical cancer (OR$_{adj}$ = 2.55; 95% CI: 1.63, 3.98). Among males, HPV 16/18 positivity was borderline significantly associated only with lifetime history of cancer (OR$_{adj}$ = 1.81; 95% CI: 0.98, 3.33).

**DISCUSSION**

We examined the associations between HPV serum infection and history of cancer among a large population sample of U.S. adults aged 18–59 years. HPV infection with two of the high-risk strains (HPV 16 or HPV 18) was associated with lifetime history of cancer, history of eight selected cancers, and lung cancer in the study population as a whole and among females, as well as with cervical

### Table 1. Weighted characteristics of the study population—National Health and Nutrition Examination Survey (NHANES) 2003–2010

| Characteristic    | Number | Weighted percent | s.e. |
|------------------|--------|------------------|------|
| **Sex**          |        |                  |      |
| Males            | 6180   | 49.98            | 0.39 |
| Females          | 6579   | 50.02            | 0.39 |
| **Age**          |        |                  |      |
| 18–29 years      | 3341   | 23.72            | 0.58 |
| 30–39 years      | 3264   | 24.66            | 0.56 |
| 40–49 years      | 3331   | 27.66            | 0.55 |
| 50–59 years      | 2823   | 23.96            | 0.64 |
| **Race**         |        |                  |      |
| White            | 5853   | 67.99            | 1.77 |
| Black            | 2602   | 11.52            | 0.88 |
| Others           | 4304   | 20.49            | 1.39 |
| **Education**    |        |                  |      |
| < high school    | 3231   | 16.71            | 0.66 |
| = high school    | 3041   | 24.03            | 0.61 |
| > high school    | 6478   | 59.26            | 0.94 |
| **Poverty income ratio** | | | |
| $< 1$            | 3444   | 18.68            | 0.67 |
| 1 $\leq$ PIR $\leq$ median | 4060 | 26.46 | 0.71 |
| $> $median       | 5255   | 54.86            | 0.97 |
| **Body mass index** | | | |
| $\geq 25$ kg/m$^2$ | 8743 | 65.95 | 0.69 |
| **Smoking**      |        |                  |      |
| Yes              | 5904   | 48.11            | 0.86 |
| **Alcohol use**  |        |                  |      |
| Yes              | 8690   | 78.01            | 0.84 |
| **Lifetime cancer** | | | |
| Yes              | 508    | 4.88             | 0.26 |
| **Selected eight cancers** | | | |
| Yes              | 177    | 1.44             | 0.16 |
| **Cervical cancer** | | | |
| Yes              | 124    | 0.97             | 0.12 |
| **Oesophageal cancer** | | | |
| Yes              | 2      | 0.02             | 0.02 |
| **Larynx/windpipe cancer** | | | |
| Yes              | 2      | 0.04             | 0.03 |
| **Lung cancer**  |        |                  |      |
| Yes              | 11     | 0.09             | 0.03 |
| **Mouth/tongue/lip cancer** | | | |
| Yes              | 5      | 0.05             | 0.02 |
| **Ovarian cancer** | | | |
| Yes              | 20     | 0.13             | 0.04 |
cancer among females. These associations remained significant after adjusting for some possible confounders, such as age, race, smoking, and drinking status.

The finding of significant positive associations between HPV 16/18 seropositivity and cervical cancer is not surprising, given the well-established aetiologic role that HPV 16 and 18 play in the development of cervical cancer (Clifford et al., 2003; Bosch et al., 2008). National surveillance data have shown higher incidence of cervical cancer among blacks (Watson et al., 2008); similarly, in the present analysis, non-Hispanic blacks had the highest HPV 16/18 seropositivity. A similar trend has also been reported in a previous analysis of NHANES (2003–2004) participants aged 14–59 (Markowitz et al., 2009). However, our data suggest that race was not a significant predictor of the association between HPV seropositivity and history of cervical cancer in the multivariable model.

### Table 2. Prevalence of seropositivity of human papillomavirus (HPV) types 16 or 18 according to cancer history (National Health and Nutrition Examination Survey (NHANES) 2003–2010)

| Variable | HPV 16/18 | Weighted percent HPV 16/18 + | s.e. |
|----------|-----------|-----------------------------|------|
| **Lifetime history** | | |
| Cancer    | +/-       | 130/378 22.8 | 1.8  |
| Non-cancer | +/-       | 1757/10494 13.6 | 0.3  |
| **Eight selected** | | |
| Cancer    | +/-       | 66/111 22.8 | 0.3  |
| Non-cancer | +/-       | 1821/10761 13.6 | 0.3  |
| **Cervix** | | |
| Cancer    | +/-       | 52/72 22.8 | 0.3  |
| Non-cancer | +/-       | 1835/10800 13.6 | 0.3  |
| **Oesophagus** | | |
| Cancer    | +/-       | 0/2 22.8 | 0.3  |
| Non-cancer | +/-       | 1887/10870 13.6 | 0.3  |
| **Larynx** | | |
| Cancer    | +/-       | 1/3 22.8 | 0.3  |
| Non-cancer | +/-       | 1886/10871 13.6 | 0.3  |
| **Lung** | | |
| Cancer    | +/-       | 5/6 22.8 | 0.3  |
| Non-cancer | +/-       | 1882/10866 13.6 | 0.3  |
| **Mouth** | | |
| Cancer    | +/-       | 1/4 22.8 | 0.3  |
| Non-cancer | +/-       | 1886/10868 13.6 | 0.3  |
| **Ovary** | | |
| Cancer    | +/-       | 6/14 22.8 | 0.3  |
| Non-cancer | +/-       | 1881/10858 13.6 | 0.3  |
| **Prostate** | | |
| Cancer    | +/-       | 1/13 22.8 | 0.3  |
| Non-cancer | +/-       | 1886/10859 13.6 | 0.3  |
| **Rectum** | | |
| Cancer    | +/-       | 1/0 22.8 | 0.3  |
| Non-cancer | +/-       | 1886/10872 13.6 | 0.3  |
The finding of significant associations between HPV 16/18 seropositivity and lifetime history of cancer adds epidemiological evidence to the carcinogenicity potential of HPV 16 and 18 on other tissues in addition to anogenital and oropharyngeal sites. The observed HPV–cancer association was also consistent with the high HPV 16/18 sero-prevalence seen in this study among participants with a history of cancer in comparison with those without any cancer.

The results of this study represent one of the most comprehensive pictures of U.S. population-based HPV infection in the context of cancer. The data consisted of four NHANES surveys between 2003 and 2010. In addition, both males and females were included in the analysis. A major limitation is the cross-sectional nature of the survey, which prevented the examination of temporal associations between HPV infection and cancers, and the fact that only antibodies to L1, the HPV structural proteins, were assessed in NHANES. It was not possible to determine the time frame between HPV testing and cancer occurrence. It was also not clear whether the observed seropositivity represented recent, latent, or reactivation of infection, and to what extent the natural history of seroreactivity differed between controls and cancer cases, whose immune response may be compromised due to cancer treatment. While HPV capsid antibody levels are stable during several years of follow-up (af Geijerstam et al, 1998) and HPV 16/18 had relatively persistent antibody responses, although with some variations (Carter et al, 2000; Ho et al, 2004), the use of HPV seropositivity in aetiological cancer studies is still emerging and need to be confirmed in tumour tissues. Other limitations include the rudimentary assessment of smoking history which is the main risk factor for many cancer sites, and the omission of other potential confounders such as lifetime number of sex partners, which has been shown to be associated with a increased risk of HPV infection (Markowitz et al, 2009) and is likely associated with smoking status as well. The relatively small number of cases for some cancers, and the fact that the outcomes were based on self-reported history, and for selected common cancers only, may have contributed to the wide error bars and confidence intervals observed. In addition, for some lethal cancer types, such as ovarian or lung cancer, a personal history of cancer likely captures only those who survived, thus introducing a survivorship bias. This bias, together with recall bias, may have affected the relative magnitude of the associations observed; for example, the association between HPV infection and history of lung cancer is unusually stronger than the association HPV–cervical cancer. It is possible that lung cancer survival differs according to the presence of HPV infection, as is well established for head and neck cancer, thus selecting for survival only those cases that are HPV positive.

However, our results suggest an association in the U.S. general population between seropositivity of high-risk HPV types (16/18) and lifetime history of cancer, cancer of the cervix and lung, and history of any of eight cancers that were selected because of their potential association with HPV. Further investigations into these associations using tumour tissues, a larger sample size, and a prospective design are warranted. With increasing coverage of HPV vaccine in the U.S., future NHANES data and sample collection may allow further detailed evaluation of the HPV–cancer relationship, as well as the population impact of HPV vaccination on cancer prevention.

### Table 3. Associations between seropositivity of human papillomavirus (HPV) types 16 or 18 and selected history of cancer outcomes

| Cancer       | Adjusted odds ratioa | 95% confidence interval | Sexb |
|--------------|----------------------|-------------------------|------|
| Lifetime     |                      |                         |      |
|              | 1.68                 | 1.35–2.09               | F + M|
|              | 1.81                 | 0.98–3.33               | M    |
|              | 1.57                 | 1.22–2.01               | F    |
| Eight selected |                    |                         |      |
|              | 2.63                 | 1.78–3.90               | F + M|
|              | 2.57                 | 0.67–9.83               | M    |
|              | 2.57                 | 1.64–4.03               | F    |
| Cervix       |                      |                         |      |
|              | 2.55                 | 1.63–3.98               | F    |
| Lung         |                      |                         |      |
|              | 5.14                 | 1.29–20.44              | F + M|
|              | 5.04                 | 1.34–18.92              | F    |

*Models were adjusted for sex, age, education, race, poverty income ratio (PIR), body mass index (BMI), smoking status, and alcohol-use status. Sampling weights were applied to account for the complex sampling National Health and Nutrition Examination Survey (NHANES) design.

b*Models were run on the whole population with both females and males (F + M), and for male-M or female-F only data sets.

---

**REFERENCES**

af Geijerstam S, Kibir M, Wang Z, Koskela P, Pukkala E, Schiller J, Lehtinen M, Dillner J (1998) Stability over time of serum antibody levels to human papillomavirus type 16. *J Infect Dis* 177(6): 1710–1714.

Alam S, Conway MJ, Chen HS, Meyers C (2008) The cigarette smoke carcinogen benzo[a]pyrene enhances human papillomavirus synthesis. *J Virol* 82(2): 1053–1058.
Anantharaman D, Gheit T, Waterboer T, Hale G, Carreira C, Abeli-Andrekan B, McKay-Chopin S, Zaridze D, Muxika A, Szeszenia-Dabrowska N, Lissowska J, Mates D, Janout V, Foretova L, Bencko V, Rudnai P, Fabianova E, Tjonneland A, Travis RC, Boeing H, Quiros JR, Johansson M, Krogh V, Bueno-de-Mesquita HB, Kotaneniu A, Clavel-Chapelon F, Weiderpass E, Johansson M, Pawlita M, Scelo G, Tommasino M, Brennan P (2014) No causal association identified for human papillomavirus infections in lung cancer. *Cancer Res* 74(13): 3525–3534.

Bosch FX, Burchell AN, Schiffman M, Giuliano AR, de Sanjose S, Bruni L, Tartofolo-Luna G, Kjaer SK, Munoz N (2008) Epidemiology and natural history of human papillomavirus infections and type-specific implications in cervical neoplasia. *Vaccine* 26(Suppl 1): K1–16.

Carter JJ, Koutsky LA, Hughes JP, Lee SK, Kuypers J, Kiviat N, Galloway DA (2003) About BMI for Adults. Available at http://www.bjcn.org/healthweight/assessing/bmi/adult_bmi/index.html.

Chaturvedi AK, Engels EA, Pfeiffer RM, Hernandez JG, Xiao W, Kim E, Jiang B, CDC Centers for Disease Control and Prevention (2011) About BMI for Adults. Available at http://www.cdc.gov/healthyweight/assessing/bmi/Adult_bmi/index.html.

Clifford GM, Smith JS, Plummer M, Munoz N, Franceschi S (2003) Human papillomavirus types in invasive cervical cancer worldwide: a meta-analysis. *Br J Cancer* 88(1): 63–73.

Das D, Van Doren J, Schlottmann S, Kelly P, Puchalski D, Ruiz W, Boerckel P, Kessler J, Antonello JM, Green T, Brown M, Smith J, Chirmule N, Barr E, Jansen KU, Esser MT (2005) Optimization and validation of a multiplexed luminesc assay to quantify antibodies to neutralizing epitopes on HPV antibodies compared with HPV DNA in cervical sera. *Int J Cancer* 68(1): 54–59.

Clifford GM, Smith JS, Plummer M, Munoz N, Franceschi S (2003) Human papillomavirus infections in invasive cervical cancer worldwide: a meta-analysis. *Br J Cancer* 88(1): 63–73.

Dillner J, Knest J, Boman J, Lehtinen M, Aarfelt R, Pukkala E, Schiller J, Maetela J, Aroma A (1998) Sero-epidemiological association between human-papillomavirus infection and risk of prostate cancer. *Int J Cancer* 75(4): 561–567.

Dillner J, Knest J, Boman J, Lehtinen M, Aarfelt R, Pukkala E, Schiller J, Maetela J, Aroma A (1998) Sero-epidemiological association between human-papillomavirus infection and risk of prostate cancer. *Int J Cancer* 75(4): 561–567.

Dillner J, Knest J, Boman J, Lehtinen M, Aarfelt R, Pukkala E, Schiller J, Maetela J, Aroma A (1998) Sero-epidemiological association between human-papillomavirus infection and risk of prostate cancer. *Int J Cancer* 75(4): 561–567.

Dillner J, Knest J, Boman J, Lehtinen M, Aarfelt R, Pukkala E, Schiller J, Maetela J, Aroma A (1998) Sero-epidemiological association between human-papillomavirus infection and risk of prostate cancer. *Int J Cancer* 75(4): 561–567.

Dillner J, Knest J, Boman J, Lehtinen M, Aarfelt R, Pukkala E, Schiller J, Maetela J, Aroma A (1998) Sero-epidemiological association between human-papillomavirus infection and risk of prostate cancer. *Int J Cancer* 75(4): 561–567.