|  |  |
| --- | --- |
| Association between HPV Vaccination and Insurance Status in the united states  National Health and Nutrition Examination Survey 2013-2014 | Abstract  **Background**: Human papillomavirus is a highly prevalent, asymptomatic sexually transmitted disease that is strongly linked to most oral and genital cancers. Health insurance and cost has been a major focus for federal financing programs to improve vaccination rates. **Objectives**: To determine if health insurance and HPV vaccinations are associated and if age, race, education, and ratio of family income to poverty modifies this association. **Methods**: Using NHANES 2013-2014 survey data, find weighted descriptive statistics of the study population, complete chi-square tests to determine significance of variables, and finally complete a multivariate logistic regression examining health insurance and HPV vaccination. **Results**: Those who were vaccinated for HPV were statistically significantly younger, had some level of college education or more, and covered with health insurance. When controlling for confounding variables, the logistic regression found that those without health insurance were 52% less likely to be vaccination for HPV. **Conclusions**: Having health insurance is a key component to HPV vaccinations in the United States.  Anjana Renganathan  Advanced Data Analytics (Spring 2020) |

**Introduction**

Human papillomavirus (HPV) is a group of viruses that cause one of the most prevalent sexually transmitted infections in the world. Cases are asymptomatic and subside almost entirely after a year or two. However, a few strains of the virus have subtle effects that can linger in the background for decades, quietly accumulating in the form of genital warts or precancerous lesions. Type 6 and 11 of HPV strains cause genital warts, while types 16,18, 31, 33, 45, 52, and 58 are oncogenic. The lesions from oncogenic strains of HPV are strongly tied to oral and most genital cancers (Senkomago et al., 2019). In fact, the majority of cases of HPV-associated cancers have a link to HPV infection from a maximum of 99.7% for cervical cancers to a minimum of 63% for penile cancers(Walboomers et al., 1999; Watson et al., 2008).

This is especially concerning when combined with the fact that 79 million Americans are currently infected with HPV, with an additional 14 million new infections per year. It is predicted that HPV is so common most sexually active people will get at least one form of it(Watson et al., 2008). Between the subtlety of symptoms, the high prevalence across the nation, and the severity of the consequences, HPV has been a critical, but difficult to address problem for society.

Therefore, Gardasil, an HPV vaccine developed in 2006, was a breakthrough preventative measure for such a prevalent, subtle disease(The Henry J. Kaiser Family Foundation, 2018). It covered 4 strains of HPV and was initially only offered to women age 9 to 26. A second-generation vaccine, Gardasil 9, was approved in 2014 and covered 9 strains of HPV. Approval for use was gradually expanded out to both men and women from age 9 to 45 in October 2018(The Henry J. Kaiser Family Foundation, 2018). However, current recommendations focus on children under the age of 11 or 12, as this is before initiation of any sexual activity. The HPV vaccine, like most other vaccines, only prevents infection against HPV strains the person has not contracted. However, the same studies that justified the FDA expanding the age range of use found that the vaccination was 88% effective in preventing the warts, precancerous lesions and cancers caused by HPV for women ages 26 to 45(The Henry J. Kaiser Family Foundation, 2018). Current recommendations by the Advisory Committee on Immunization Practices advise that children ages 11 to 14 receive two doses over 6 months, while those from 15 to 26 years of age receive three doses over six months(The Henry J. Kaiser Family Foundation, 2018).

Current up-to-date(UTD) HPV vaccination coverage in US adolescents’ hovers around 50%, with 69% of girls and 63% of boys having received at least one dose(Neubrand, Radecki Breitkopf, Rupp, Breitkopf, & Rosenthal, 2009). This disparity in coverage is most likely attributable to the dose-series nature of the shot, with two to three doses occurring over 6 months. Barriers to vaccine dose completion include family poverty, gender, race/ethnicity, type of insurance, reasons for clinic visits, urbanicity, and state(Lu, O’Halloran, & Williams, 2015; Neubrand et al., 2009; The Henry J. Kaiser Family Foundation, 2018; Walker et al., 2019).

The impact of health insurance status is one of the most intriguing. In a study using 2012 National Health Interview data, researchers found that HPV vaccination coverage was two to three times higher for insured adults when compared to uninsured adults(Lu et al., 2015). This disparity was larger in those that did not have a regular physician, and those who had fewer physician contacts per month(Lu et al., 2015).

This difference in vaccination coverage and completion has been noted by the government and dealt with four different public financing efforts, including Medicaid, the Vaccines for Children Program, the Immunization Grant Program and CHIP(Sanders Thompson, Arnold, & Notaro, 2012; The Henry J. Kaiser Family Foundation, 2018). These programs typically cover all costs and target underserved or at-risk populations, including Medicaid-eligible adults and children, American Indian or Alaskan Native populations, the uninsured, the underinsured, and individual state’s CHIP-eligible children(The Henry J. Kaiser Family Foundation, 2018). Between the drop in vaccine dosages from two to three, extension of the program to men and women up to the age of 45 and full coverage between private and public providers, vaccination initiation and UTD coverage is set up to hopefully, steadily increase over the coming years with consequential long-term reductions in HPV-associated cancer rates.

In this study, we aim to reaffirm and examine the association between insurance status and HPV vaccination initiation using national Health and Nutrition Examination Survey data from 2013 to 2014. We will also determine whether age, race, education level and family-income-to-poverty ratio modified the association between insurance status and HPV vaccination initiation.

**Methods**

The data used was pulled from the National Health and Nutrition Examination Survey (NHANES) for the 2013 to 2014 survey cycle. This nationally representative population health survey is completed biennially by the National Center for Health Statistics (NCHS), an extension of the Centers for Disease Control and Prevention (CDC) in the United States. It involves a complex survey design involving oversampling, survey non-response and post-stratification that is described elsewhere(Johnson et al., n.d.). Additional modules are provided describing proper use and analysis of the data (Centers for Disease Control and Prevention & National Center for Health Statistics, 2019). Since this dataset is publicly available and deidentified, IRB approval was not required.

The study population included all adults above the age of 20 that participated in the Questionnaire components of NHANES and provided answers for all variables in question. There was no exclusion or inclusion criteria outside of the unintentional limits by the survey process and in the creation of a complete case set. The use of the education level variable meant that only participants older than 20 were included in the analysis. The final survey sample included 3,448 participants.

The variables collected included race/ethnicity, age, ratio of family income to poverty, education level, self-reported HPV vaccination status, and health insurance status. Race/ethnicity (RIDRETH3) had seven levels, including Mexican Americans, Other Hispanics, Non-Hispanic Whites, Non-Hispanic Blacks, Non-Hispanic Asians, and Others, which included the Multi-Racial category. Age (RIDAGEYR) was collected in years at the time of the survey. Ratio of family income to poverty (INDFMPIR) was a continuous variable from 0 to more than 5. This was used over family income as a measure of financial stability because it accounts for the number of family members and differences in state poverty level. Education level (DMDEDUC2) had 5 levels including less than 9th grade, 9th to 11th grade, high school graduate or equivalent, some college or an associate degree, and 4-year college graduate or more. The HPV vaccination data as collected by NHANES was divided by gender, so IMQ040 covered female participants, while IMQ070 covered male participants. This was united into a single variable on self-reported HPV vaccination status. Finally, health insurance (HIQ010) described whether the individual had health insurance or not.

A directly acyclic graph (DAG) showing the causal diagram was constructed with the use of ‘DAGitty’ (Textor, Hardt, & Knüppel, 2011). Two versions were created to show versions of the causal chain before and after controlling for certain variables. Statistical analysis was completed using R Studio (version 1.1.463), a publicly available statistical coding program, along with use of the packages ‘RNHANES’, ‘survey’, ‘dplyer’, ‘tidyverse’, ‘jtools’, ‘dataMaid’ and ‘expss’ (Demin & Jeworutzki, n.d.; Jacob & Long, 2020; Lumley, 2020; Petersen & Ekstrøm, 2019; True, 2016; Wickham et al., 2019). Data management included ‘passing over’ NA or missing variables due to the complex survey design and recoding the gender segregate HPV vaccination variables into a single united variable. A table of weighted descriptive statistics for the study population was completed, with chi-square tests to determine if there was a difference between those with and without the HPV vaccination for each of the variables. Then, a multivariate logistic regression was performed for HPV vaccination and health insurance after controlling for age, race, education level and income level. The reference group was non-Hispanic white high school graduates with health insurance. A Wald test was included for each variable to determine which significantly added to the model. Finally, the multivariate logistic regression was tested for assumptions including no multicollinearity and linearity.

For reproducibility, the Coding2Share toolkit was fully utilized, with the statistical analysis and codebook made available in the author’s GitHub, as seen in Appendix A.

**Results**

Two DAGs were constructed to show the causal framework between health insurance status and HPV vaccinations, as shown in Figures 1a and 1b. They include confounders like race/ethnicity, age, education, and poverty, as well as ancestors of the outcome like marital status, sexual education, vaccine cost, having a regular healthcare provider, and HPV awareness. The first model shows the unchanged model with confounders, while the second shows the DAG model for the logistic regression, after the confounders are controlled for.

A picture containing text, map

Description automatically generatedA close up of a map

Description automatically generated

**Figure 1a and 1b**. Directed acyclic graphs showing the causal framework between health insurance status and self-reported HPV vaccination status for this study.

The weighted sociodemographic characteristics of the study population by HPV vaccination status are shown in Table 1. Chi-square tests found that there was a statistically significant difference in age, education level, and health insurance status between the two populations. Mean age of those without HPV vaccination was a little less than twice as much as those with HPV vaccinations, at 40.8 and 27.2 years respectively (x2 (39) =554.8, p<0.0001). When looking at education levels, most of the group with HPV vaccinations had at least some college, while those without were more evenly dispersed across education levels (x2 (4) =24.7, p<0.0001). Finally, the vaccination status of those with health insurance statistically significantly differed from those without, at 83.7% and 77.5% respectively (x2 (1) =5.7, p=0.0303).

**Table 1.** Weighted sociodemographic characteristics of NHANES 2013-2014 study population, stratified by HPV vaccination status.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | Have you been vaccinated for HPV? | |
|  |  | Yes | No |
|  |  | n=276 | n=3172 |
| Age (mean ± SD) | | | |
|  | | 27.2±8.1 | 40.8±11.2 |
| Race/Ethnicity (%) | | | |
|  | Mexican | 7.7% | 10.1% |
|  | Other Hispanic | 5.9% | 5.9% |
|  | White | 61.7% | 63.0% |
|  | African American | 14.2% | 11.9% |
|  | Asian | 5.9% | 5.9% |
|  | Other, including Multi-Racial | 4.6% | 3.2% |
| Education Level (%) | | | |
|  | <9th grade | 1.9% | 3.7% |
|  | 9th to 11th grade | 8.3% | 10.2% |
|  | High School Graduate or GED | 17.1% | 22.6% |
|  | Some College or Associates Degree | 45.6% | 31.5% |
|  | 4-year College Graduate or more | 27.1% | 32.0% |
| Health Insurance (%) | | | |
|  | Yes | 83.7% | 77.5% |
|  | No | 16.3% | 22.4% |
| Family-Income-to-Poverty Ratio (median [IQR]) | | 1.0 [2.1-4.42] | 1.27[2.9-5] |

The multivariate logistic regression found that, after controlling for age, race/ethnicity, education level, and family-income-to-poverty ratio, those without health insurance had 52% lower odds of HPV vaccination when compared to those that had health insurance (OR: 0.48, 95% CI: 0.35-0.66, p=0.02). Wald tests found that age, and health insurance were significant to the regression model (Age: F (1,3) =110.0, p=0.002; Health Insurance: F (1,3) =19.9, p=0.021).

When checking assumptions for the logistic regression, binary dependent variable, independence of observations and adequate sample size were all met. When checking multicollinearity with a VIF cutoff at 10, there was evidence of multicollinearity for race and education. Both continuous variables, age, and ratio of family income to poverty, were linear.

**Discussion**

This study found that those who self-report HPV vaccinations were more likely to be younger, have at least some level of college education and have health insurance. Age is generally consistent with the situation, but it and all the other variables are not consistent across the literature. HPV vaccine initiation had varying associations with age, race/ethnicity, education level or family income across studies, as found in a systematic review of 25 studies (Kessels et al., 2012). The same review found that the group of studies reporting on the association of health insurance and vaccination rates was split evenly between the conclusion that children with health insurance had higher odds of HPV vaccination and that there was no association at all(Kessels et al., 2012). Another, more extensive study on health insurance and HPV vaccination found significant differences between those without insurance, those with public insurance, and those with private insurance. 43.4% of women with private insurance had been vaccinated for HPV, compared to 30.4% of those with public insurance and 20.9% of those without health insurance(Lu et al., 2015).

Conclusions from this study, as well as most of the literature, should be taken in consideration of the limitations of the situation. The HPV vaccination is still relatively new and recommended usage by the FDA has changed at least 4 times since it was released in 2006 (The Henry J. Kaiser Family Foundation, 2018). Additionally, much of the focus for HPV vaccinations is for age groups below 20, which were not captured in our dataset. As such, we did not capture the entire target population, but rather the earlier cohorts of children who received the vaccination at a younger age in early years of release and aged into the study population at 2013-2014 as well as older teens and adults who received ‘catch-up’ vaccinations. Also, since many of the variables were self-reported, they are most likely subject to recall and self-serving biases. The study is also limited by the data it is based on. We were limited to 20 and older since responses on health insurance were only collected for that age range. Finally, it’s important to remember that NHANES, while keenly representative on demographics, is not a geographically representative survey(Johnson et al., n.d.). In 2013-2014, they only sampled about 16 PSUs. And since both Medicaid eligibility and HPV rates differ significantly across states, it is important to realize that the scope of the survey’s sampling points could not even cover all 50 states.

Future research is endless, with the situation around HPV vaccinations still evolving and the initial cohort of patients aging into years where HPV-associated cancers might present themselves. The relationships between health insurance and HPV vaccination initiation and completion should be more definitively and comprehensively studied in one project, as current analyses are scattered through different papers with varying definitions. It would also be interesting to see a longitudinal study on the change in HPV vaccinations that is nuanced to changes in FDA recommendations. Research useful to health care providers and public health workers would focus on the differential barriers and limitations of targeted subpopulations. Initial qualitative research has shown high concern on the cost of vaccines without insurance, paired with low awareness of public financing projects of programs like Vaccines for Children that would cover all costs of the program(The Henry J. Kaiser Family Foundation, 2018).

**References**

Centers for Disease Control and Prevention, & National Center for Health Statistics. (2019). NHANES Tutorials. Retrieved from https://wwwn.cdc.gov/nchs/nhanes/tutorials/default.aspx

Demin, G., & Jeworutzki, S. (n.d.). Package expss. Retrieved May 2, 2020, from https://cran.r-project.org/web/packages/expss/index.html

Jacob, A., & Long, A. (2020). *Package ‘ jtools .’*

Johnson, C. L., Paulose-Ram, R., Ogden, C. L., Carroll, M. D., Kruszan-Moran, D., Dohrmann, S. M., & Curtin, L. R. (n.d.). National health and nutrition examination survey. Analytic guidelines, 1999-2010 (N. C. for H. S. (U.S.), Ed.). *Vital and Health Statistics. Series 10, Data from the National Health Survey*. Retrieved from https://stacks.cdc.gov/view/cdc/21305

Kessels, S. J. M., Marshall, H. S., Watson, M., Braunack-Mayer, A. J., Reuzel, R., & Tooher, R. L. (2012). Factors associated with HPV vaccine uptake in teenage girls: A systematic review. *Vaccine*, *30*(24), 3546–3556. https://doi.org/10.1016/j.vaccine.2012.03.063

Lu, P. J., O’Halloran, A., & Williams, W. W. (2015). Impact of health insurance status on vaccination coverage among adult populations. *American Journal of Preventive Medicine*, *48*(6), 647–661. https://doi.org/10.1016/j.amepre.2014.12.008

Lumley, A. T. (2020). *Package ‘ survey .’*

Neubrand, T. P. L., Radecki Breitkopf, C., Rupp, R., Breitkopf, D., & Rosenthal, S. L. (2009). Factors associated with completion of the human papillomavirus vaccine series. *Clinical Pediatrics*, *48*(9), 966–969. https://doi.org/10.1177/0009922809337534

Petersen, A. H., & Ekstrøm, C. T. (2019). dataMaid : Your Assistant for Documenting Supervised Data Quality Screening in R. *Journal of Statistical Software*, *90*(6). https://doi.org/10.18637/jss.v090.i06

Sanders Thompson, V. L., Arnold, L. D., & Notaro, S. R. (2012). African American Parents’ HPV Vaccination Intent and Concerns. *Journal of Health Care for the Poor and Underserved*, *23*(1), 290–301. https://doi.org/10.1353/hpu.2012.0007

Senkomago, V., Henley, S. J., Thomas, C. C., Mix, J. M., Markowitz, L. E., & Saraiya, M. (2019). *Human Papillomavirus – Attributable Cancers — United States , 2012 – 2016*. *68*(33), 2012–2016.

Textor, J., Hardt, J., & Knüppel, S. (2011). DAGitty. *Epidemiology*, *22*(5), 745. https://doi.org/10.1097/EDE.0b013e318225c2be

The Henry J. Kaiser Family Foundation. (2018). *The HPV Vaccine : Access and Use in the U. S.* (October), 1–6. Retrieved from www.kff.org

True, L. (2016). *Package ‘ RNHANES .’*

Walboomers, J., Jacobs, M., Manos, M., Bosch, F., Kummer, J., Shah, K., … Munoz, N. (1999). Human Papillomavirus Is a Necessary Cause. *Journal of Pathology*, *189*(May), 12–19.

Walker, T. Y., Elam-Evans, L. D., Yankey, D., Markowitz, L. E., Williams, C. L., Fredua, B., … Stokley, S. (2019). National, regional, state, and selected local area vaccination coverage among adolescents aged 13-17 years - United States, 2018. *Morbidity and Mortality Weekly Report*, *68*(33), 718–723. https://doi.org/10.15585/MMWR.MM6833A2

Watson, M., Saraiya, M., Ahmed, F., Cardinez, C. J., Reichman, M. E., Weir, H. K., & Richards, T. B. (2008). Using population-based cancer registry data to assess the burden of human papillomavirus-associated cancers in the United States: Overview of methods. *Cancer*, *113*(10 SUPPL.), 2841–2854. https://doi.org/10.1002/cncr.23758

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., … Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

**Appendix A.**

Link to Author’s Github and Project Repository: https://github.com/arenganathan28/MPH\_ADA\_FinalProject