

THE CORRELATION BETWEEN HPV VACCINATION RATE AND INCOME INEQUALITY

Zhifei Xie

Garrison Forest School, Owings Mills, Maryland, USA

ABSTRACT

According to the 2018 National Immunization Survey - Teen (NIS-Teen), the human papillomavirus (HPV) vaccination coverage in the U.S. increased from 48.6% to 51.1%.[1] Certain factors contribute to disparity between teenagers receiving HPV vaccination.[1]The factors are geography, race, gender, education level, household income, and etc. Within these factors, household income and income inequality were chosen to be the focus of this study. The relationship between HPV vaccination coverage in the U.S., the household income of interviewed individuals, and the Gini index in the U.S. have been studied in RStudio [2]. By merging the NIS-Teen data of vaccination rate and Gini index data in RStudio, charts and graphs are formed to illustrate the relationship between HPV vaccination rate and income inequality. There seem to be limited correlations between vaccination rate and Gini index, but unexpected connections between vaccination rate and household income have been found.

KEYWORDS

HPV vaccination, Gini index, income inequality, RStudio, American adolescents

1. INTRODUCTION

Among all the sexually transmitted infections, the human papillomavirus (HPV) is the most prevalent one.[3] The likelihood of HPV infection is more than 80% for women and 90% for men.[4] In the United States, 79 million people are currently infected with HPV and approximately 14 million become newly infected each year.[5] Generally, HPV goes away by itself, but when it doesn't, it can cause health problems like genital warts, cervical cancer, and other cancers.[5] Fortunately, the HPV vaccines can effectively protect against several types of cancers and other diseases caused by the virus.[6] The HPV vaccine is given with a series of shots, but the dosage depends on the ages of the individuals. Children who are nine to fourteen years old only need two doses, and people aged from fifteen to twenty-six should receive three doses.[7] There are three types of HPV vaccines that prevent infections: Gardasil, Gardasil-9, and Cervarix. Cervarix (the bivalent HPV vaccine) prevents infection with high-risks HPV types 16 and 18 that cause 70% of cervical cancers. Gardasil (the quadrivalent HPV vaccine) has improved on Cervarix in that it also protects against HPV types 6 and 11, the two that causes 90% of genital warts. And in 2016 Gardasil-9 (the nonavalent HPV vaccine) replaced the other two since it prevents the four types covered by Cervarix and Gardasil, as well as additional five types (31, 33, 45, 52, and 58) that account for 10 to 20% of cervical cancers. Currently, Gardasil-9 is the only HPV vaccine in use in the U.S. [8].According to Centers for Disease Control and Prevention (CDC), there are about 34,800 cases of cancer caused by HPV each year in the U.S., and the 9valent HPV vaccine (Gardasil-9) can efficiently prevent 92% of them.[5]

However, the HPV vaccination rate among teenagers is disconcerting even when the effectiveness of the vaccine is unquestionable. Walker et al. mentioned in their report that "uptake of the HPV vaccine remains modest" and that "rates of HPV vaccination have been

lower than desired by public health advocacy groups to achieve strong herd immunity in the US, particularly among the age groups in which it is recommended (11–12 years old).”[1] This phenomenon and uneven vaccination coverage are caused by numerous factors, and one of them could be the affordability of insurances and access to health care providers for people with different levels of income.

From a research that invested on the association between parental income and HPV vaccine uptake for 12 years old girls in Norway, it is shown that high maternal income has a positive association with high probability of initiating HPV vaccination.[9] A similar situation could happen in the U.S., especially when the income inequality becomes severer than before. From 1983 to 2016, the share of U.S. aggregate wealth for the upper-income households increased from 60% to 79%; but on the other hand, the median-income households held about a half of what they had in 1983 and the low-income families had only 4% of the aggerate wealth, indicating that the wealth gap is on a rise.[10] Moreover, the Gini index of the U.S. increased from 0.43 in 1990 to 0.49 in 2018, which means that the income inequality gets worse and the high-income individuals are receiving a larger portion of the total income of the population than before, as shown in figure 1.[11] This increasing income inequality may influence the HPV vaccination rate in that high-income individuals are more likely than their low- or median-income counterparts to devote resources to their children’s health care, which thereby causing uneven vaccination coverage among adolescents. Thus, the Gini index or household income level may be able to explain the decreasing prevalence of HPV vaccination and the disparity of vaccination rates among different groups of individuals.

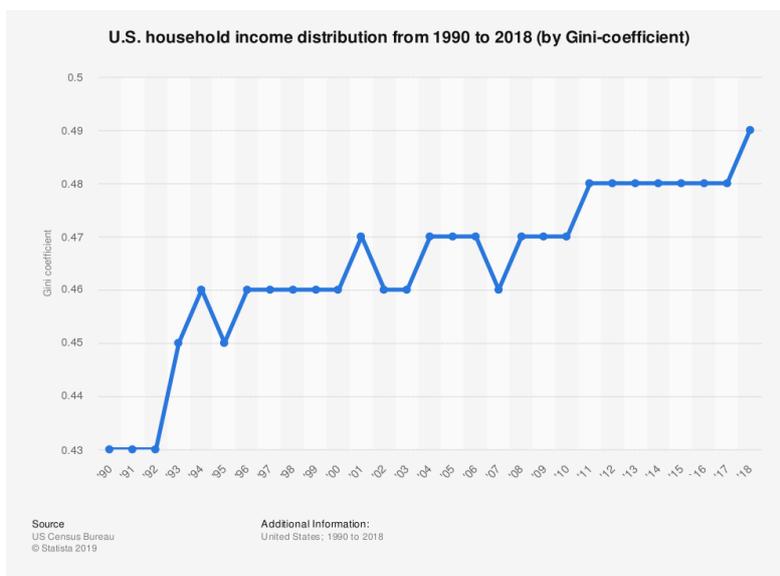


Figure 1. U.S. household income distribution from 1990 to 2018 (by Gini-coefficient) [11]

2. COLLECTING DATA AND INFORMATION

The National Immunization Survey - Teen (NIS-Teen) is a survey conducted by CDC to monitor the annual vaccine coverage of 13 to 17 years old adolescents in the 50 states, the District of Columbia, selected local areas, and U.S. territories. Random-digit–dialed samples of cell phone numbers are used to conduct interviews with parents or guardians of eligible teenagers.[1]The content of the interview includes vaccination history, teen and household health conditions, demographic and socioeconomic questions, vaccination provider information, and health insurance information.[1] The vaccine providers identified in the interviews are then mailed

questionnaires that require them to specify the teen’s vaccination history, which are then used to estimate the actual vaccination coverage (in other words, the estimated vaccination coverage is based on the health care providers’ information but not based on the information provided in the phone calls).[12] In 2018, the NIS-Teen survey conducted complete household interviews for 38,706 teens and received adequate provider data for 18,700 teens.[12] The survey also recorded the household incomes of the selected teens and categorized them into four groups with regard to poverty level: income > \$75,000 (at or above poverty); income <= \$75,000 (at or above poverty); below poverty; and unknown.[12]

In this research, the results from the 2018 NIS-Teen survey are used and all of the data from the 2018 NIS-Teen public-use data file is imported into RStudio where the HPV vaccination rate in each state is calculated. The formula for this calculation is the total number of teens who have received a total of three HPV vaccine shots within each state divided by the total number of teens interviewed in that state. Another set of data relating to the income levels is formed by a similar calculation: the total number of teens who have received a total of three HPV vaccine shots within each income level over the total number of teens interviewed within that income level. Then, the data about the Gini index in the U.S. is also imported into RStudio. The Population Reference Bureau [13]provides the Gini index of each U.S. state during two intervals: 2006-2010 and 2013- 2017. Since the HPV vaccination data from the NIS-Teen is from 2018, the most recent 2013-2017 Gini index in each state is picked to merge with vaccination data. The two data frames - the HPV vaccination rate and the Gini index - are joined by states, and a scatter plot (figure 2) below representing the data in each state is formed. In the plot, the dark red dots represent the HPV vaccination rates and the steel blue dots represent the Gini indexes.

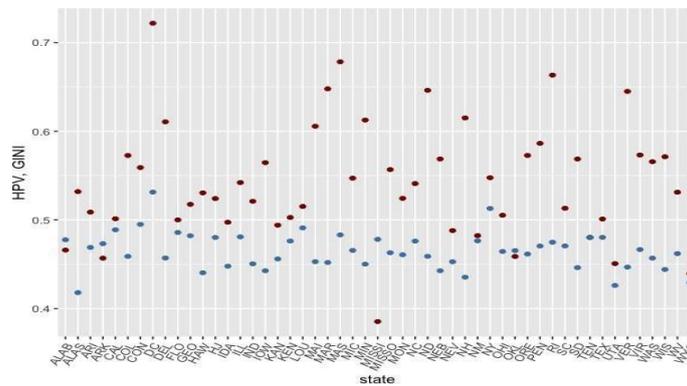


Figure 2. HPV vaccination rates (red) and Gini Coefficients (blue) in 50 states

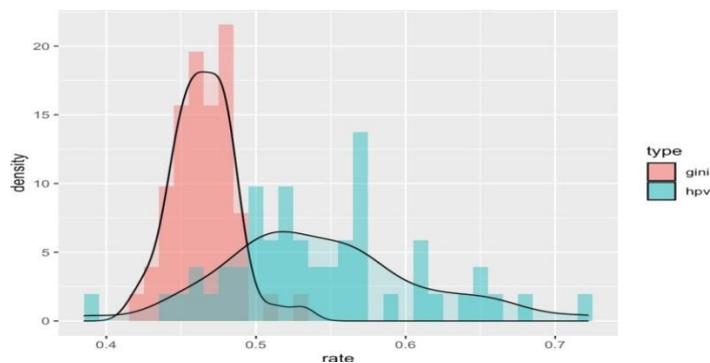


Figure 3. HPV vaccination rate (blue) and Gini index (red) histogram and their kernel density plot (black)

3. THE CORRELATION BETWEEN HPV VACCINATION RATE AND GINI INDEX

To better analyse the relationship between the HPV vaccination rate and the Gini coefficient, a histogram (figure 3) was also created. The “rate” on the x-axis represents either the HPV vaccination rate (blue) or Gini index (red), and “density” on the y-axis represents probability density (in other words, the count of data which falls into that range represented by different bins on the x-axis). The red columns in this graph show that most of the states have their Gini indexes laying in the range between 0.4 and 0.5. The density of the Gini index peaks at the range of 0.47 to 0.48, giving that more than 20 states have Gini indexes within this scope. The blue columns illustrate that the HPV vaccination rate of different states lay in the range of 0.4 to 0.7 and mostly cluster between 0.5 and 0.6. And the density of HPV vaccination rate peaks when the vaccination rate is within the scope of 0.56 and 0.575, giving that roughly 14 states have an HPV vaccination rate within this range. Similar to the scatter point plot (figure 2), this histogram (figure 3) also shows that there is a limited correlation between the Gini index and the HPV vaccination rate. The histogram generally shows the distribution or spread of the data, so if two variables are linearly and positively correlated, the shape and spread of their graph should overlap; however, such overlap doesn’t occur in figure 3.

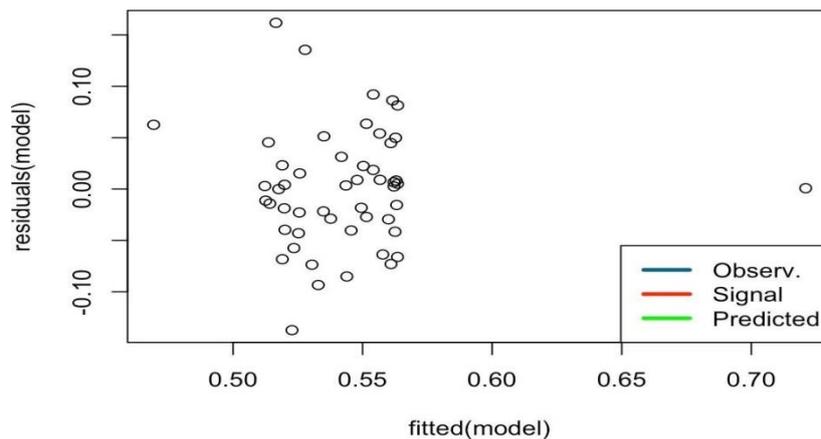


Figure 4. Residual plot for the regression fits of the Gini index and vaccination rate

In this residual plot of the model (figure 4), a regression curve that has cubic polynomials is used. “Fitted” on the x-axis represents the ideal model or equation that shows the mathematical relationship between Gini index and vaccination rate. “Residuals” on the y-axis represent the difference between the fitted model and the actual relation between Gini index and vaccination rate. The method is repeated several times with different polynomial degrees (polynomial to the 1st, 2nd, and 4th), and none of them yields a regression curve that would result in an equation representing the relationship between the two variables. Therefore, it is concluded that the HPV vaccination rate and the Gini index have a weak connection.

4. THE CORRELATION BETWEEN HPV VACCINATION RATE AND HOUSEHOLD INCOME

Since there seems to be a limited correlation between the HPV vaccination rate and the Gini index, the focus is then switched to the connection between HPV vaccination and household income. The 2018 NIS-Teen public-use data file categorized the household income of the interviewed family into four groups (income > \$75,000; income <= \$75,000; below poverty; and

unknown). And the data file also includes the income to poverty ratio of the households interviewed. Using the existing income status data in the data file and the previously calculated HPV vaccination rate within each income level, an income level plot (figure 5) with regard to HPV vaccination rate is created.

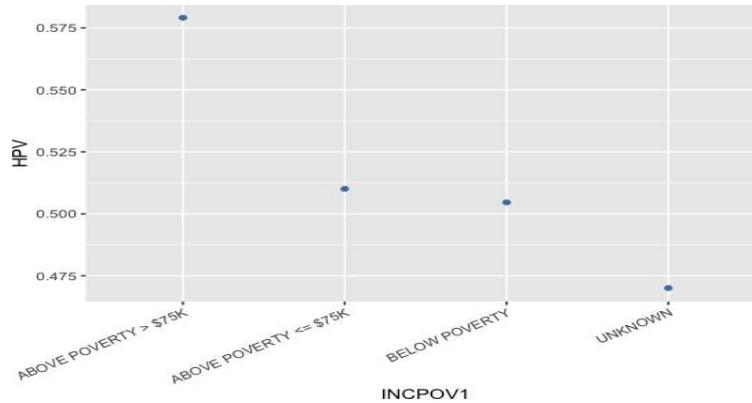


Figure 5. HPV vaccination rate at different income level

In figure 5, “INCPOV1” on the x-axis represents the poverty status of the households and the “HPV” on the y-axis represents the HPV vaccination rate within each income level. It is shown on the graph that the households earning incomes higher than \$75,000 have the highest HPV vaccination rate among the four poverty statuses. The HPV vaccination rate is more than 57.5% in this range of income, which is relatively higher than that of the other income levels. Interestingly, the HPV vaccination rate of the households earning lower or equal to \$75,000 is roughly 51.25%, which is really close to the 50.6% vaccination rate of the households earning incomes below the poverty line. And lastly, the unknown group accounts for the vaccination rate for households that did not provide adequate income information. It is expected to see that the households earning the most among the four groups are more likely to have their children receiving HPV vaccines. However, it is rather unexpected that the households earning a relatively “median” income (the households that have lower or equal \$75,000) among the four poverty statuses have a similar likelihood to vaccinate the children with the households earning incomes below the poverty line.

In figure 5, “INCPOV1” on the x-axis represents the poverty status of the households and the “HPV” on the y-axis represents the HPV vaccination rate within each income level. It is shown on the graph that the households earning incomes higher than \$75,000 have the highest HPV vaccination rate among the four poverty statuses. The HPV vaccination rate is more than 57.5% in this range of income, which is relatively higher than that of the other income levels. Interestingly, the HPV vaccination rate of the households earning lower or equal to \$75,000 is roughly 51.25%, which is really close to the 50.6% vaccination rate of the households earning incomes below the poverty line. And lastly, the unknown group accounts for the vaccination rate for households that did not provide adequate income information. It is expected to see that the households earning the most among the four groups are more likely to have their children receiving HPV vaccines. However, it is rather unexpected that the households earning a relatively “median” income (the households that have lower or equal \$75,000) among the four poverty statuses have a similar likelihood to vaccinate their children with the households earning incomes below the poverty line.

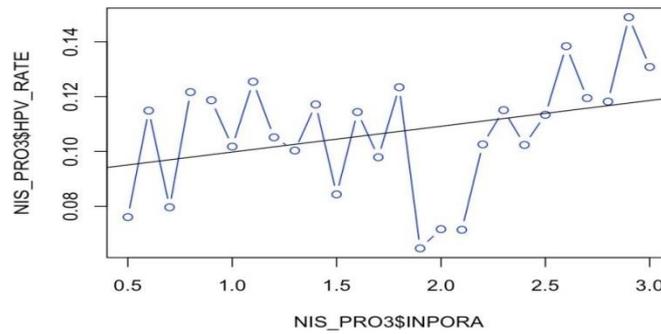


Figure 6. Three doses HPV vaccination rate in different income-to-poverty ratio household

Similar results are also seen from the graph above (figure 6). In this graph, “NIS_PRO3\$INPORA” on the x-axis means the ratio of the family’s income to the poverty level. Since it is the ratio between household income and poverty level (income: poverty line), the larger the number on the x-axis the higher the income of the family.

“NIS_PRO3\$HPV_RATE” on the y-axis shows the three doses HPV vaccination rate. From the graph, the three doses HPV vaccination rate is the highest when the income to poverty ratio is in the range between 2.5 and 3.0, showing that it is way more likely for teenagers to get vaccinated if their families’ incomes are relatively high. The three doses vaccination rate is the lowest when the income to poverty ratio is roughly in the range between 1.8 to 2.2, which similar to the result gathered from figure 5, shows that the median income earners are less likely to vaccinate their children than expected. The possible reasons for high-income households having high vaccination rates are straightforward: they have the money to pay for high-quality health care providers; they are generally better educated; they live in better communities; and etc. However, the reasons behind median income households having relatively low vaccination rates are not very clear.

Similar results are also seen from the graph above (figure 6). In this graph, “NIS_PRO3\$INPORA” on the x-axis means the ratio of the family’s income to the poverty level. Since it is the ratio between household income and poverty level (income: poverty line), the larger the number on the x-axis the higher the income of the family. “NIS_PRO3\$HPV_RATE” on the y-axis shows the three doses HPV vaccination rate. From the graph, the three doses HPV vaccination rate is the highest when the income to poverty ratio is in the range between 2.5 and 3.0, showing that it is way more likely for teenagers to get vaccinated if their families’ incomes are relatively high. The three doses vaccination rate is the lowest when the income to poverty ratio is roughly in the range between 1.8 to 2.2, which, similar to the result gathered from figure 5, shows that the median income earners are less likely to vaccinate their children than expected. The possible reasons for high-income households having high vaccination rates are straightforward: they have the money to pay for high-quality health care providers; they are generally better educated; they live in better communities; and etc. However, the reasons behind median income households having relatively low vaccination rates are not very clear.

5. CONCLUSIONS

In this study, the correlation between HPV vaccination rates among US teenagers and income inequality between households was analysed. The results derived from the statistical analysis of NIS-Teen and PRB data show that there is a limited correlation between the HPV vaccination rate and the Gini index. The plausible limitations of the study can explain such a result. First, the

data about the Gini index in the U.S. is from 2013 to 2017, but the HPV vaccination data is from 2018; the slight difference in time may cause some unwanted effects. Second, the Gini index represents income inequality of a whole state, but the HPV vaccination rate is generated only from the interviewed households that have adequate provider data, which is not representative of the vaccination rates for the teens as a whole within each state. In another comparative study that used National Immunization Survey NIS-Teen 2016 data, it was concluded that gender, race/ethnicity, religiosity, age, and maternal education are all predictors of HPV vaccination for U.S. adolescents, and that conservative and religious states are the ones with lowest HPV vaccination rate.[14] While the NIS-Teen 2018 database provided HPV vaccination data based on many factors, this research only focused on the contribution of household income to HPV vaccination rate. In order to generate more accurate results, the data about Gini index and HPV vaccination rate should be collected from the same year. Moreover, all the other factors such as gender, age, geography, and etc. should be excluded from the data, so that the final outcome would provide the direct relationship between HPV vaccination rate and income inequality without the interference of any other factors.

Although there isn't sufficient information to indicate a relationship between HPV vaccination rate and Gini index, a clear correlation between HPV vaccination rate and household income is shown in this research. Based on the data, families with relatively higher income are more likely to vaccinate their children and that the teenagers in median and low-income households are similarly less likely to receive the HPV vaccination. Many factors contribute to this phenomenon, and one that is considered in this study is that families with median income or income below poverty may tend to prioritize their spending on other life necessities instead of on insurances that cover HPV vaccine. Further, HPV vaccine is not required for school entry [1], which may lead many parents or guardians to believe that it is unnecessary for their children to receive such vaccination. In 2016, there were only three jurisdictions that require HPV vaccine for school, but even if the schools launch such requirements, a substantial number of parents may oppose to it out of different concerns.[15] Another important factor is the health care provider's recommendation for HPV vaccine. In a national study on the HPV vaccination among adolescent girls, it is shown that among the daughters who had received one or more HPV vaccine doses, 90% of the surveyed mothers reported to receive healthcare providers' recommendations.[16] It is possible that high income families with access to private health care providers often get vaccine-related recommendations, so their likelihood of being vaccinated is greater. However, even though families with income above \$75k have a higher HPV vaccination rate than the rest, that vaccination rate is only about 57.5%. Therefore, HPV vaccination is far from being prevalent in the United States, and perhaps starting to require HPV vaccination for school and educate parents on relevant topics are the first step towards change.

In this study, the correlation between HPV vaccination rates among US teenagers and income inequality between households was analysed. The results derived from the statistical analysis of NIS-Teen and PRB data show that there is a limited correlation between the HPV vaccination rate and the Gini index. Many reasons can explain such a result. First, the data about the Gini index in the U.S. is from 2013 to 2017, but the HPV vaccination data is from 2018; the slight difference in time may cause some unwanted effects. Second, the Gini index represents income inequality of a whole state, but the HPV vaccination rate is generated only from the interviewed households that have adequate provider data, which is not representative of the vaccination rates for the teens as a whole within each state. Although there isn't sufficient information to indicate a relationship between HPV vaccination rate and Gini index, a clear correlation between HPV vaccination rate and household income is shown by the graphs in this research. It seems that families with relatively higher income are more likely to vaccinate their children and that the teenagers in median and low-income households are similarly less likely to receive the HPV vaccination. Many factors contribute to this phenomenon, and one that is considered in this

study is that families with median income or income below poverty may tend to prioritize their spending on other life necessities instead of on insurances that cover HPV vaccine. Further, HPV vaccine is not required for school entry [16], which may cause many parents or guardians to believe that it is unnecessary for their children to receive such vaccination. Another important factor is the health care provider's recommendation for HPV vaccine. Other studies have shown that adolescents who were recommended for receiving HPV vaccine are more likely to be vaccinated [16]. Thus, it is possible that high-income families that possess private health care providers often get vaccine related recommendations, so their likelihood of being vaccinated is greater. However, even though families with income above \$75k have a higher HPV vaccination rate than the rest, that vaccination rate is only about 57.5%. Therefore, HPV vaccination is far from being prevalent in the United States, and perhaps starting to require HPV vaccination for school is the first step towards change.

ACKNOWLEDGMENTS

I would like to express my deep gratitude to Dr. Ying Qin Taylor, my research mentor, for her patient guidance during the planning and development of this research work. Her constructive advice and valuable suggestions help me a tremendously in using RStudio to analyse the data. Her time and assistance are greatly appreciated.

REFERENCES

- [1] T Y Walker, L D Elam-Evans, and D Yankey, (2018) "Regional, State, and Selected Local Area Vaccination Coverage Among Adolescents Aged 13-17 Years - the United States", *MMWR Morb Mortal Wkly Rep*, Vol. 68, pp. 718–723. DOI: <http://dx.doi.org/10.15585/mmwr.mm6833a2>.
- [2] RStudio Team, 2020. [Online]. Available: <http://www.rstudio.com/>.
- [3] Hirth J. (2019). Disparities in HPV vaccination rates and HPV prevalence in the United States: a review of the literature. *Human vaccines & immunotherapeutics*, 15(1), 146–155. <https://doi.org/10.1080/21645515.2018.1512453>
- [4] Satterwhite CL, Tortrone E, Meites E, Dunne EF, Mahajan R, Ocfemia MCB, Su J, Xu F, Weinstock H. Sexually transmitted infections among US women and men: prevalence and incidence estimates, 2008. *Sex Transm Dis*. 2013;40(3):187–193. doi:10.1097/OLQ.0b013e318286bb53.
- [5] "STD Facts - Human papillomavirus (HPV)," *Centers for Disease Control and Prevention*, 20Aug-2019. [Online]. Available: <https://www.cdc.gov/std/hpv/stdfact-hpv.htm>.
- [6] M. Jenco, "CDC: Teens' HPV vaccination rates improve slightly," *American Academy of Pediatrics*, 29-Aug-2020. [Online]. Available: <https://www.aappublications.org/news/2019/08/22/teenvaccination082219>.
- [7] "HPV," *Centers for Disease Control and Prevention*, 15-Aug-2019. [Online]. Available: <https://www.cdc.gov/hpv/hcp/schedules-recommendations.html>.
- [8] "Human Papillomavirus (HPV) Vaccines," *National Cancer Institute*. [Online]. Available: <https://www.cancer.gov/about-cancer/causes-prevention/risk/infectious-agents/hpv-vaccine-factsheet>.
- [9] B. Feiring, I. Laake, T. Molden, et al. "Do parental education and income matter? A nationwide register-based study on HPV vaccine uptake in the school-based immunisation programme in Norway," *BMJ Open*, 01-May-2015. [Online]. Available: <https://bmjopen.bmj.com/content/5/5/e006422.citation-tools>.
- [10] R. I. and R. K. Juliana Menasce Horowitz, "Trends in U.S. income and wealth inequality," *Pew Research Center's Social & Demographic Trends Project*, 17-Aug-2020. [Online]. Available: <https://www.pewsocialtrends.org/2020/01/09/trends-in-income-and-wealth-inequality/>.
- [11] P. by E. Duffin and S. 18, "U.S. income distribution of individual earnings 1990-2017, by Gini coefficient," *Statista*, 18-Sep-2019. [Online]. Available: <https://www.statista.com/statistics/219846/us-income-distribution-of-individual-earnings-bygini-coefficient/>.
- [12] U.S. Department of Health and Human Services (DHHS). National Center for Immunization and Respiratory Disease. The 2018 National Immunization Survey - Teen, Atlanta, GA: Centers for Disease Control and Prevention, 2020.

- [13] "Data Center: US Data," *Population Reference Bureau*. [Online]. Available: <https://www.prb.org/usdata/indicator/gini/table>.
- [14] Franco, M., Mazzucca, S., Padek, M. *et al.* Going beyond the individual: how state-level characteristics relate to HPV vaccine rates in the United States. *BMC Public Health* **19**, 246 (2019). <https://doi.org/10.1186/s12889-019-6566-y>
- [15] A. L. North and L. M. Niccolai, "Human Papillomavirus Vaccination Requirements in US Schools: Recommendations for Moving Forward," *American journal of public health*, Oct-2016. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5024362/>.
- [16] L. M. Kester, G. D. Zimet, J. D. Fortenberry, et al. "A national study of HPV vaccination of adolescent girls: rates, predictors, and reasons for non-vaccination," *Maternal and child health journal*, Jul-2013. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3719972/>.