Effect of COVID-19 burden on changes in HPV vaccination coverage among females between 2019 and 2022

Rayleen Lewis

**Authors**

* Rayleen Lewis

**Author affiliations**

1. College of Public Health, University of Georgia, Athens, GA, USA.

Corresponding author: rml20258@uga.edu

Disclaimer: The opinions expressed in this article are the author’s own and don’t reflect their employer or University.

# 1. Abstract

## 1.1 Background:

Human papillomavirus vaccines have been available since 2006, and as of 2019, 108 countries had introduced HPV vaccine into their national immunization programs. In 2020, immunization services globally experienced disruptions due to the COVID-19 pandemic. My objective was to assess the association between country-level COVID burden and changes in HPV vaccination coverage between 2019 and 2022.

## 1.2 Methods:

Data on annual coverage, cumulative number of COVID cases between Jan. 1, 2020, and Dec. 31, 2022, country income level, vaccination program characteristics, and sociodemographic information were publicly available from the World Health Organization and United Nations. Multivariable linear regression models, modeling the difference in 2022 and 2019 HPV vaccination coverage, were performed with number of cumulative COVID cases as the primary predictor; additional country vaccination program and sociodemographic characteristics were considered for adjustment. Analyses were restricted to complete cases (N = 98). LASSO regression and regression tree models were also considered. Models were compared using R2 values.

## 1.3 Results:

There was a weak, positive, linear association between the difference in HPV vaccination coverage and number of cumulative COVID cases (𝛽(se) = 0.000769 (0.00036), p-value = 0.034). Variability in the coverage difference was highest for countries with the lowest COVID burden and decreased as COVID burden increased. After adjustment for all vaccination program and sociodemographic characteristics, coverage difference and COVID burden were not associated (beta = -0.000065 (0.00064), p-value = 0.919). COVID burden was an important predictor of difference in coverage among countries with a median resident age of <26 years.

## 1.4 Conclusions:

After accounting for additional country-level characteristics, COVID burden was not associated with the change in coverage between 2019 and 2022 for countries overall. There is high variability between countries in the ability and methods used to assess both HPV vaccination coverage and COVID burden. It is unclear if the null association observed would be replicated if countries used similar methodology for assessing national health metrics.

# 2. Introduction

Human papillomavirus (HPV) is the most common sexually transmitted infection globally.(1) There are about 40 mucosal HPV types, and at least 14 are considered high-risk, or oncogenic, and may lead to cancer.(2) Globally, cervical cancer is the most common HPV-attributable cancer and is the 4th most common cancer among women. In 2022, almost 350,000 women died of cervical cancer, with over 90% of these deaths occurring among women who live in low and middle income countries.(3)

Several HPV vaccines are available to prevent infections that cause most HPV-attributable cancers. The first vaccine product, a quadrivalent vaccine that protects against 4 HPV types, entered the market in 2006; the United States was the first country to make a national HPV vaccination recommendation.(4,5) Since then, five additional vaccine products have entered the market and over 100 countries have introduced HPV vaccination into their national immunization program.(5,6)

In December 2019, cases of a severe respiratory disease increased rapidly in China. We now know this was the start of the COVID-19 pandemic.(7) Globally, this pandemic disrupted many medical services, including immunization programs.(8) Several studies have demonstrated the country-specific impact of the COVID-19 pandemic on HPV vaccination and HPV-related disease.(9–13) Additionally, a global assessment of the effect of the pandemic on HPV vaccination coverage has been performed.(8) However, these analyses focused on pre- and post-pandemic comparisons rather than considering the magnitude of COVID burden in the country. To address this gap, my objective was to assess the association between country-level COVID burden and changes in HPV vaccination coverage between 2019 and 2022. I hypothesized that decreases in HPV vaccination coverage between 2022 and 2019 were greater among countries with high burden of COVID-19 in 2020-2022. A secondary objective of this work was to develop a regression tree to assist with classifying countries by their change in HPV vaccination coverage using other country-level HPV vaccine program and sociodemographic characteristics. The purpose of this regression tree is to identify key sociodemographic indicators that may be used by policy makers for pandemic preparedness efforts to prevent reductions in HPV coverage if a new epidemic or pandemic were to affect a country.

# 3. Methods

## 3.1 Data acquisition

Data on HPV vaccination coverage and vaccination program details are available at the country level through the World Health Organization’s (WHO) [HPV Dashboard](https://www.who.int/teams/immunization-vaccines-and-biologicals/diseases/human-papillomavirus-vaccines-(HPV)/hpv-clearing-house/hpv-dashboard).(14) Briefly, coverage data reflect administrative and official HPV vaccination coverage reported annually through the WHO/UNICEF Joint Reporting Form on Immunization. The HPV Dashboard includes country income level (low, lower middle, upper middle, high), whether there was a national HPV vaccination recommendation, year of vaccine introduction, primary delivery strategy (school-based, facility based, varies by region, mixed), number of doses in routine schedule, whether the program is gender neutral (i.e., targets males and females), WHO region, and year-specific coverage from 2010 to 2023.

Country-specific weekly COVID-19 case and death counts are available starting on January 4, 2020, through the WHO [COVID-19 dashboard](https://data.who.int/dashboards/covid19/data?n=o).(15) In August 2023, WHO stopped requiring weekly reporting, although some countries continued to voluntarily report COVID-19 data.

Country-level demographics including life expectancy of females at birth and population size were available from the [United Nations](https://population.un.org/wpp/downloads?folder=Standard%20Projections&group=CSV%20format).(16)

All data sources are publicly available for download and were downloaded on January 26-27, 2025. Each data source was imported and merged based on a country code identifier, which was present in all datasets along with country name.

The data were first restricted to countries with vaccination coverage estimates in both 2019 and 2022. The outcome variable was coded as HPV vaccination coverage of at least one dose among females in 2022 minus coverage in 2019. Of the 100 countries with coverage information from both years, 2 countries were missing some vaccine program characteristic information, one of which was also missing income level. All other countries had complete data for all variables. Complete cases were used for all analyses (N = 98).

## 3.2 Country-level characteristics

COVID burden between January 4, 2020, and December 31, 2022, was defined as the cumulative number of COVID cases per 100,000 persons. Additionally, cumulative number of COVID deaths per 100,000 was also calculated.

Vaccination program characteristics include WHO region, country income level, presence of a national vaccination recommendation, year of HPV vaccination program introduction, primary vaccine delivery strategy (e.g., school-based), whether the program offers gender neutral vaccination (i.e., female-only vaccination or females and males), and recommended number of doses.

Country-level demographic characteristics include total population (thousands), population density, median age of residents, population growth rate, total fertility rate, crude death rate, life expectancy at birth, and net migration rate.

Additional details for each characteristic can be found in the Supplementary Material.

## 3.3 Statistical analysis

All data management and statistical analyses were performed using R version 4.4.2 (2024-10-31) within RStudio Version 2024.12.0+467.

### 3.3.1 Descriptive analyses

Univariate analyses of the outcome and each country-level characteristic were performed. A histogram of the difference in coverage was assessed to determine the approximate distribution of the outcome. The mean (standard deviation (sd)) and median (interquartile range (IQR)) were calculated for each continuous country characteristic (e.g., number of cumulative COVID cases). For categorical variables, the distribution (n, %) of each level was calculated. Bivariate associations between each characteristic and the change in HPV vaccination coverage between 2022 and 2019 were visually assessed using scatterplots (with a LOESS line of best fit for visualization purposes only) for continuous variables and violin plots for categorical variables.

### 3.3.2 Multivariable modeling

Multivariable linear regression was performed using the lm function to assess the association between the difference in HPV vaccination coverage and cumulative number of COVID cases, a measure of COVID burden. An unadjusted model and fully adjusted model (adjusted for all country level demographic and vaccination program characteristics) were performed. Additionally, models adjusted for each country level characteristic individually were performed. For each model, the effect of COVID case count on the difference in HPV vaccination coverage between 2019-2022 (beta, standard error, p-value from t-test with a null hypothesis of beta = 0) and model performance metrics (R-squared, AIC, BIC) were extracted and compared. Selection of a final model was based on subject matter expertise, which model had the highest R-squared, and which model had the lowest AIC and/or BIC. Model parameters were presented for the final selected model. Because there are many potential predictors, LASSO regression was also performed to reduce the risk of overfitting. LASSO regression was fit for a grid of penalty values from 1E-5 to 1E2 on a log scale using the glmnet engine and tidymodels framework. The highest tuning value without a reduction in R2 was selected as the final tuning value. Beta estimates and R2 values from the final multivariable linear regression and LASSO regression models were compared. Diagnostic plots (e.g., predicted vs observed) for both models were visually inspected to assess model fit.

### 3.3.3 Regression tree analysis

Regression tree analyses were conducted using the rpart package. This package builds a regression tree using a two step process. First, the variable that best splits the data into two groups is identified. The process is repeated on the two resulting groups. This is repeated until subgroups either reach the minimum sample size or the model cannot be improved by additional variable splits. Second, cross validation is used to trim leaves from the tree to reduce the risk of overfitting. All characteristics were included as potential predictors in the regression tree analysis. R2 was estimated for the best fitting model for comparison with the multivariable linear and LASSO regressions.

# 4. Results

## 4.1 Descriptive analyses

### 4.1.1 Univariate analyses

#### 4.1.1.1 Difference in HPV vaccination coverage between 2022 and 2019

The difference in HPV vaccination coverage between 2022 and 2019 appears normally distributed centered around 0 (i.e., no change in coverage).([Figure 1](#fig-Figure1)) The mean (sd) difference in coverage was -5.5 (23.4); the median (IQR) was 0.0 (-15.0, 6.0).([Table 1](#tbl-Table1_coverage_covid))

|  |
| --- |
| Figure 1: Histogram of difference in HPV vaccination coverage between 2022 and 2019 |

|  |
| --- |
| Table 1: HPV vaccination coverage and COVID burden among countries with HPV vaccination programs in 2019 and 2022 |

#### 4.1.1.2 COVID burden

Between January 1, 2020, and December 31, 2022, the mean (sd) number of cases per 100,000 persons was 8,132.1 (6,470.6). The median (IQR) was lower at 6,960.7 (2,387.3, 13,567.2).([Table 1](#tbl-Table1_coverage_covid)) The mean 3-year COVID death rate was 56.9 (47.6) deaths per 100,000 persons; the median was similar at 51.1 (16.7, 82.8).

#### 4.1.1.3 Vaccination program characteristics

The median year of vaccine introduction was 2014 (2010, 2017).([Table 2](#tbl-table1_vac_program)) School-based delivery programs were the most common (60 / 100 (60%)). A majority of programs had a gender nuetral vaccination recommendation and a 2-dose routine recommendation (rather than a 1-dose recommendation).

|  |
| --- |
| Table 2: HPV vaccination program characteristics among countries with HPV vaccination programs in 2019 and 2022 |

#### 4.1.1.4 Sociodemographic characteristics

Sociodemographic metrics were presented in [Table 3](#tbl-table1_demo). The most common WHO region was the Euro region, and the most common country income level was high-income. Mean life expectancy was 75.2 (6.5) years.

|  |
| --- |
| Table 3: Sociodemographic characteristics among countries with HPV vaccination programs in 2019 and 2022 |

### 4.1.2 Bivariate analyses

#### 4.1.2.1 Bivariate associations between the difference in HPV vaccination coverage between 2022 and 2019 and COVID burden

There was a weak linear association between cumulative COVID cases and the difference in HPV vaccination coverage between 2019 and 2022 ([Figure 2](#fig-Figure2)). Of note, there is a lot of variation in the coverage difference among countries with fewer cumulative COVID cases per 100,000 persons.

|  |
| --- |
| Figure 2: Scatterplot of difference in HPV vaccination coverage between 2019 and 2022 and cumulative COVID cases per 100,000 |

A similar association was seen when assessing cumulative COVID deaths (Supplementary material). Because these variables are both measures of COVID burden and they exhibited similar patterns, the remainder of the analysis focused on COVID case burden.

#### 4.1.2.2 Bivariate associations between the each country demographic and vaccination program characteristic and the difference in HPV vaccination coverage between 2022 and 2019

The distribution of the difference in vaccination coverage by each country-level demographic variable can be seen in [Figure 3](#fig-Figure4). For each WHO region, the distribution of difference in coverage centers around 0, with the SEAR and EUR showing a more narrow distribution. Similarly, the distribution of coverage difference was centered around 0 regardless of country income level, and high-income countries had the tightest distribution. The patterns in the continuous variables are less straightforward, particularly as there appear to be potential outliers. Based on the figure below, total population may have a linear, positive association with coverage difference, while life expectancy and median age show a negative quadratic association and fertility rate may show a positive quadratic association.

|  |
| --- |
| Figure 3: Distribution of difference in HPV vaccination coverage between 2022 and 2019 by each country demographic variable |

There were some differences in the distribution of coverage difference by vaccine program characteristics ([Figure 4](#fig-Figure5)). Countries with a facility-based delivery strategy, gender neutral vaccination program, and 2-dose recommendation had distributions of difference in coverage with a smaller portion of countries in the distribution region indicating a decrease in coverage (i.e., most of the data was centered around zero or was greater than 0). There was not a strong association with difference in coverage and year of implementation.

|  |
| --- |
| Figure 4: Associations between vaccine program characteristics and difference in HPV vaccination coverage between 2022 and 2019 and cumulative COVID deaths per 100,000 |

## 4.2 Multivariable modeling

### 4.2.1 Model building and selection

#### 4.2.1.1 Multivariable linear regression

For the unadjusted and all adjusted multivariable linear regression models, COVID burden beta estimates, standard errors, p-values for the effect of cumulative COVID cases on the difference in coverage, and performance metrics (R2, AIC, BIC) are presented in [Table 4](#tbl-modelcomp). Each performance metric has been shaded using a sliding color scale; a light color indicates worse performance, darker color indicates better performance, and the “best” performing model is indicated using bold text. The fully adjusted model (adjusted for all country-level demographic and vaccination program characteristics) had the highest R2. The model adjusted for median resident age had the lowest AIC and BIC. The association between change in coverage and COVID burden were similar in the fully adjusted model (𝛽=-0.000065) and model only adjusted for median age (𝛽=-0.000182) ([Table 4](#tbl-modelcomp), see supplementary material for all model parameters for median age adjusted model). Preliminary modeling will focus on the fully adjusted model. However, none of these models fit particularly well based on the low R2 values and diagnostic plots in [Figure 5](#fig-lm-diagnostics). Based on the predicted vs observed plot, many observations with relatively high or low observed coverage difference values are further from the line of perfect fit. Additionally, there appears to be a slight pattern in the residuals; relatively small predicted values are more likely to have negative residuals and moderate values are more likely to have large positive values. Given that the beta is switching directions (positive/negative) based on the adjustment variables and is very close to zero ([Table 4](#tbl-modelcomp)), there does not appear to be a strong association between COVID burden and difference in coverage regardless of adjustment.

|  |
| --- |
| Table 4: Preliminary model comparison for models assessing the effect of cumulative COVID cases on the difference in HPV vaccination coverage |

|  |
| --- |
| Figure 5: Diagnostic plots for the fully adjusted multivariable linear regression model: predicted vs observed values (left panel) and predicted values vs residuals (right panel) |

#### 4.2.1.2 LASSO regression

Because of the large number of coefficients in the full model, LASSO regression model was also performed for regularization. A grid of penalty values from 1E-5 to 1E2 on a log scale was considered for tuning. [Figure 6](#fig-Lasso_fit) shows the R2 value for each model tuning parameter in the left panel, and the residuals from the run with each tuning parameter in the right panel. Each color represents a particular run. R2 is relatively stable for tuning parameters values from 1E-5 to 1E-2; therefore, a tuning value of 1E-2 was selected for the final model.

|  |
| --- |
| Figure 6: LASSO regression fit diagnostic plots showing the effect of tuning parameter on model R2 (left panel) and residuals by predicted values for each tuning parameter (right panel) |

#### 4.2.1.3 Regression tree analysis

Median age was the variable identified for the first node, splitting the countries based on median age < 26 years. For countries with a median age < 26, countries were further split by cumulative COVID cases per 100,000 persons (>=155, <155). Of these, countries with a higher COVID burden (n = 17) had a 27 percentage point decrease in HPV vaccination coverage on average; countries with a lower COVID burden (n = 7) had an average increase in coverage of 1.9 percentage points. For countries with a median age >=26 years, further dichotomizations were made based on an additional median age dichotomization, population density, fertility rate, migration rate, and number of HPV doses recommended.

|  |
| --- |
| Figure 7: Regression tree of coverage difference between 2022 and 2019 using all predictors |

### 4.2.2 Model results and comparison

Estimates from the fully adjusted model are shown in [Table 5](#tbl-fully_adj). There was not a significant association between cumulative number of COVID cases and the difference in HPV vaccination coverage between 2019-2022 using the linear model (𝛽(se) = -0.00006547213 (0.0006436804), *p*-value = 0.9192541). After adjusting for all other characteristics, the difference in coverage was significantly higher and positive (indicating increase in coverage between 2019 and 2022) for all WHO regions, except the Eastern Mediterranean Region, compared to the African Region. Conclusions from the LASSO regression model were similar. However, some beta estimates other than COVID burden changed magnitude and/or direction (e.g., population growth rate, low income) compared to the linear regression model.

|  |
| --- |
| Table 5: Model parameters for fully adjusted model assessing the effect of cumulative COVID cases on the difference in HPV vaccination coverage |

Based on the regression tree ([Figure 7](#fig-tree)), median age of residents was the most important characteristic for grouping countries based on their coverage differences between 2022 and 2019. COVID burden was an important predictor among countries with a median age of < 26 years for persons residing in the country.

Of the four models considered, the regression tree had the largest R2 value (0.361). This model is also potentially the most intuitive, and the utility of estimating COVID burden is apparent in this model.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 6: R2 comparison of models   | Model | R-squared | | --- | --- | | Full multivariable linear regression | 0.344 | | Median age adjusted linear regression | 0.156 | | LASSO regression | 0.335 | | Regression tree | 0.361 | |

# 5. Discussion

Overall, there was not a significant association between COVID burden and changes in HPV vaccination coverage after adjusting for other country level characteristics. Previous COVID research found countries with high levels of internal income inequality were disproportionately impacted by COVID.(17) The observed lack of an association in the present analysis may be due to masked regional or local differences since only national level data are available.

While all countries use the same case reporting form to submit HPV vaccination coverage estimates to WHO, the methods used to generate the initial estimates vary widely. WHO does an independent assessment of each country’s data to better understand if changes in estimates from year to year may be due to methodological issues in data collection or if estimates have truly changed.(18) This may be particularly problematic if reporting and COVID burden were influenced by an additional factor, such as country income level. To mitigate potential confounding effects, country-level demographic characteristics (e.g., median age of residents), income level, and vaccination program characteristics were considered for adjustment. However, residual confounding may be possible.

A strength of this work is the development of the regression tree analysis that may be used by countries if faced with a pandemic in the future. Among countries with a younger median age of its residents, countries who are experiencing a higher burden of disease due to a pandemic may choose to prioritize resources to minimize changes in vaccination coverage. Based on the results in the present study, this may be less of a priority for countries with lower pandemic disease burden or countries with older residents on average.

This analysis is subject to several limitations. First, as discussed above, countries vary in the methods used to estimate national health indicators, including vaccination coverage. Second, the absolute number of countries was relatively small. However, these countries represent more than half of the countries reporting to the World Health Organization. Third, the analysis was restricted to countries who had HPV vaccination coverage estimates in 2019 and 2022, meaning only countries that had introduced HPV vaccination by 2019 were included. In the event of a future pandemic, it is unclear if these findings can be generalized to countries that introduced HPV vaccination after 2019 or choose not to introduce a national HPV vaccination program.

## 5.1 Conclusion

During the first two years of the COVID-19 pandemic, many countries experienced a decrease in HPV vaccination coverage. Overall, COVID burden was not a strong predictor of changes in HPV vaccination coverage. However, for countries with a younger median age of residents, higher COVID burden was associated with larger decreases in HPV vaccination coverage, highlighting the importance of developing country-specific pandemic preparedness efforts.

# 6. References

1. Sanjosé S de, Diaz M, Castellsagué X, et al. Worldwide prevalence and genotype distribution of cervical human papillomavirus DNA in women with normal cytology: A meta-analysis. *The Lancet Infectious Diseases* [electronic article]. 2007;7(7):453–459. (<https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(07)70158-5/fulltext>)

2. Lewis RM, Laprise J-F, Gargano JW, et al. [Estimated prevalence and incidence of disease-associated human papillomavirus types among 15- to 59-year-olds in the united states](https://doi.org/10.1097/OLQ.0000000000001356). *Sexually Transmitted Diseases*. 2021;48(4):273–277.

3. Organization WH. Global strategy to accelerate the elimination of cervical cancer as a public health problem. 1st ed. World Health Organization; 2020.

4. Markowitz LE, Gee J, Chesson H, et al. [Ten years of human papillomavirus vaccination in the united states](https://doi.org/10.1016/j.acap.2017.09.014). *Academic Pediatrics*. 2018;18(2S):S3–S10.

5. Bruni L, Saura-Lázaro A, Montoliu A, et al. HPV vaccination introduction worldwide and WHO and UNICEF estimates of national HPV immunization coverage 2010–2019. *Preventive Medicine* [electronic article]. 2021;144:106399. (<https://www.sciencedirect.com/science/article/pii/S0091743520304308>)

6. WHO. Human papillomavirus vaccines: WHO position paper (2022 update). *Weekly epidemiological record*. 2022;50(97):645–672.

7. Cucinotta D, Vanelli M. WHO declares COVID-19 a pandemic. *Acta Bio Medica : Atenei Parmensis* [electronic article]. 2020;91(1):157–160. (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7569573/>)

8. Casey RM, Akaba H, Hyde TB, et al. Covid-19 pandemic and equity of global human papillomavirus vaccination: Descriptive study of world health organization-unicef vaccination coverage estimates. *BMJ Medicine* [electronic article]. 2024;3(1):e000726. (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10826539/>)

9. Ivanuš U, Jerman T, Gašper Oblak U, et al. [The impact of the COVID-19 pandemic on organised cervical cancer screening: The first results of the slovenian cervical screening programme and registry](https://doi.org/10.1016/j.lanepe.2021.100101). *The Lancet Regional Health. Europe*. 2021;5:100101.

10. Gountas I, Favre-Bulle A, Saxena K, et al. Impact of the COVID-19 pandemic on HPV vaccinations in switzerland and greece: Road to recovery. *Vaccines* [electronic article]. 2023;11(2):258. (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9964352/>)

11. Silva TMRD, Nogueira de Sá ACMG, Beinner MA, et al. Impact of the COVID-19 pandemic on human papillomavirus vaccination in brazil. *International Journal of Public Health* [electronic article]. 2022;67:1604224. (<https://www.ssph-journal.org/journals/international-journal-of-public-health/articles/10.3389/ijph.2022.1604224/full>)

12. Rao SR, Kampan N, Chew KT, et al. The impact of the COVID-19 pandemic on the national HPV immunization program in malaysia. *Frontiers in Public Health* [electronic article]. 2022;10. (<https://www.frontiersin.org/journals/public-health/articles/10.3389/fpubh.2022.907720/full>)

13. Daniels V, Saxena K, Roberts C, et al. [Impact of reduced human papillomavirus vaccination coverage rates due to COVID-19 in the united states: A model based analysis](https://doi.org/10.1016/j.vaccine.2021.04.003). *Vaccine*. 2021;39(20):2731–2735.

14. WHO. Immunization, vaccines, and biologicals: HPV dashboard. (<https://www.who.int/teams/immunization-vaccines-and-biologicals/diseases/human-papillomavirus-vaccines-(HPV)/hpv-clearing-house/hpv-dashboard>)

15. WHO. WHO COVID-19 dashboard.

16. United Nations,. World population prospects 2024.

17. Su D, Alshehri K, Pagán JA. Income inequality and the disease burden of COVID-19: Survival analysis of data from 74 countries. *Preventive Medicine Reports* [electronic article]. 2022;27:101828. (<https://www.sciencedirect.com/science/article/pii/S2211335522001358>)

18. WHO. WHO UNICEF immunization coverage estimates 2023 revision. 2024;(<https://www.who.int/docs/default-source/immunization/immunization-coverage/wuenic_notes.pdf?sfvrsn=88ff590d_6#:~:text=Additional%20data%20sources:,and%20global%20vaccination%20coverage%20figures.>)