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. Summary/Abstract

# 2. Introduction

## 2.1 General Background Information

### 2.1.1 Description of data and data source

The difference in HPV vaccination coverage between 2022 and 2019 will be modeled using multivariable regression techniques, including multivariable linear regression *other information on new modeling techniques will be added*. Cumulative number of COVID cases per 100,000 persons as of December 31, 2022 was the explanatory variable. HPV vaccination program characteristics, country income information, and country-level demographics will be considered for adjustment. HPV vaccination coverage, HPV vaccine program characteristics, and cumulative number of COVID cases were extracted from publicly available data from the World Health Organization. Country sociodemographic information was publicly available from the United Nations. All information was collected at the country level, and data were restricted to countries that had HPV vaccination coverage estimates for both 2019 and 2022.

### 2.1.2 Questions/Hypotheses to be addressed

The objective of this work is to determine if changes in HPV vaccination coverage were correlated with COVID burden in 2020-2022. I hypothesize that HPV vaccination coverage was lower in 2022 than 2019 among countries with high burden of COVID-19 in 2020-2022 and comparable in countries with a low COVID-19 burden.

*This is to be determined based on methods learned later in the course, but this is an initial thought on how to implement one of the methods covered in a later module.* An additional objective of this work is to develop a classification tree analysis 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 classification 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 an epidemic or pandemic affects a country.

# 3. Methods

## 3.1 Data aquisition

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). Briefly, coverage data reflect administrative and official HPV vaccination coverage reported annually through the WHO/UNICEF Joint Reporting Form on Immunization. HPV Dashboard include country income level (low, lower middle, upper middle, high), whether there was a national HPV vaccine schedule, 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 (target males and females), WHO region, and year-specific coverage from 2010 to 2023.

Country-specific weekly COVID-19 case and death counts are available from starting in 1/4/2020 through the WHO [COVID-19 dashboard](https://data.who.int/dashboards/covid19/data?n=o). 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)

All data sources are publicly available for download and were downloaded on January 26-27, 2025.

## 3.2 Data import and cleaning

Each data source imported and merged based on country; all datasets had variables for a country code and name. Individual datasets were restricted to the variables of interest as follows:

* HPV vaccination coverage
  + HPV vaccination coverage in 2022
  + HPV vaccination coverage in 2019
* HPV vaccination program characteristics
  + WHO region
  + World Bank country income level
  + Indicator for whether HPV vaccine is offered through a national recommendation
  + Year of HPV vaccine program introduction
  + Primary HPV vaccine delivery strategy (e.g., school-based)
  + Indicator for gender neutral vaccination (i.e., female-only vaccination or females and males)
  + Recommended number of doses (1 or 2)
* United Nations country demographic data
  + Total Population, as of 1 July (thousands)
  + Population Density, as of 1 July (persons per square km)
  + Median Age, as of 1 July (years)
  + Population Growth Rate (percentage)
  + Total Fertility Rate (live births per woman)
  + Crude Death Rate (deaths per 1,000 population)
  + Life Expectancy at Birth, both sexes (years)
  + Net Migration Rate (per 1,000 population)
* COVID burden through Dec. 31, 2022
  + Cumulative number of COVID cases per 100,000 persons
  + Cumulative number of COVID deaths per 100,000 persons

The final project dataset was restricted to countries with vaccination coverage estimates in both 2019 and 2022 (N = 100). 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 included countries, 2 countries were missing some vaccine program characteristic information and 1 country was missing income level. All other countries had complete data for all variables.

## 3.3 Statistical analysis

### 3.3.1 Exploratory/Descriptive analyses

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. A histogram of the difference in coverage was graphed. Bivariate associations between each characteristic and the change in HPV vaccination coverage between 2022 and 2019 were visually assessed using scatterplots for continuous variables and violin plots for categorical variables.

### 3.3.2 Multivariable modeling

Multivariable linear regression was performed to assess the association between 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 were estimated. For each model, the effect of COVID case count on 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.

### 3.3.3 Classification tree analysis

All data management and analysis was conducted using R (Version 2024.12.0+467). Classification tree analyses were conducted using the XXXX package.

# 4. Results

## 4.1 Exploratory/Descriptive analysis

### 4.1.1 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 2019 and 2022 |

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

### 4.1.2 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 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 6,960.7 (2,387.3, 13,567.2).

### 4.1.3 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 |

### 4.1.4 Sociodempgraphic 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).

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

## 4.2 Basic statistical analysis

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

There appears to be 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 lower numbers of cumulative COVID cases.

|  |
| --- |
| 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 ([Figure 3](#fig-Figure3)). Because these variables are both measures of COVID burden and they exhibit similar patterns, the remainder of the analysis is focused on COVID case burden.

|  |
| --- |
| Figure 3: Scatterplot of difference in HPV vaccination coverage between 2019 and 2022 and cumulative COVID deaths per 100,000 |

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

The distribution of the difference in vaccination coverage by each country level demographic variable can be seen in [Figure 4](#fig-Figure4). For each WHO region, the distribution of change 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 negative quadratic association.

|  |
| --- |
| Figure 4: Distribution of difference in HPV vaccination coverage between 2019 and 2022 by each country demographic variable |

There were some differences in the distribution of coverage difference by vaccine program characteristics ([Figure 5](#fig-Figure5)). Countries with a facility-based delivery strategy, gender neutral vaccination program, and 2-dose recommendation had distributions of coverage difference 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 association.

|  |
| --- |
| Figure 5: Associations between vaccine program characteristics and difference in HPV vaccination coverage between 2019 and 2022 and cumulative COVID deaths per 100,000 |

## 4.3 Preliminary modeling

As a preliminary modeling step, beta estimates, standard errors, and p-values for the effect of cumulative COVID cases on the difference in coverage 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 and lowest AIC. The model adjusted for country income level had the lowest BIC. Therefore, preliminary modeling will focus on the fully adjusted model. However, none of these models are a particularly great fitting model, and given that the beta is switching directions (positive/negative) and very close to zero, 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 |

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. 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.

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

## 4.4 Full analysis

# 5. Discussion

## 5.1 Summary and Interpretation

## 5.2 Strengths and Limitations

## 5.3 Conclusions

# 6. References