**Title Page**

**Title**: Human Papillomavirus Vaccination Trends and Patterns Study on the Long Island

**Abbreviated Title**: Trends and Patterns of HPV study among teenagers (Around 9 to 15 year-old) on the Long Island, 2012-2021

**Authors:**

**Corresponding Author**

Fusheng Wang PhD

Department of Biomedical Informatics

Department of Computer Science

Stony Brook University

2313D Computer Science

Stony Brook, NY 11794-8330

Tel: (631)632-2594

Email: fusheng.wang@stonybrook.edu

**Abstract**

**Objectives:** This paper aims to study the temporal trends and geographical patterns of human papillomavirus (HPV) vaccine administration rate and the demographic and regional disparities by analyzing large scale immunization registry vaccine data for Long Island (LI).

**Design:** We conducted temporal and spatial analysis on the data retrieved from the New York State Immunization Information System (NYSIIS) which includes the patient’s vaccination visits spanning the years 2012 to 2023.

**Participants:** 410778 patients and 1756134 visit records were included in our dataset.

**Results:** In our study, the total HPV rate on the Long Island demonstrates a growth from 2012 to 2019 by 284.43% and slight falls in 2020 and 2021 by around 5%. Both female and male’s rate experience a similar trend. Among all the HPV records we have, around 66% patients received their first dose before 15 years old and 23.74% of them follows the recommended routine to receive 2 doses in half year.

**Conclusions**: The findings outlined by this paper emphasize the rise of HPV vaccine usage on the Long Island among teenagers from the 2012 to 2019. The results also indicate the decrease of vaccination rate among teenagers led by the COVID-19 pandemic around 2020.

**Introduction**

HPV infections cause approximately 33700 cases of cancer every year in the United States, including cervical, vaginal, penile, anal and head and neck cancers.[Cite, usHPVtypes] The Human Papillomavirus (HPV) vaccine is designed to prevent infection with certain types of HPV, which can lead to above types of cancer (anal, cervical, oropharyngeal, vulvar, vaginal, and penile) and genital warts. It is one of the most effective ways to prevent HPV-related diseases. The vaccine is typically administered through a series of injections. In the United States, the HPV vaccine has been widely available and recommended for adolescents and young adults 1. Various campaigns and initiatives have been launched to increase awareness and uptake of the vaccine. In New York, like in many other states, efforts have been made to ensure accessibility to the vaccine through healthcare providers, clinics, and public health programs. Additionally, school-based vaccination programs and outreach efforts have been utilized to reach adolescents.

This section will introduce our research background and motivation first. Then we will go through the HPV and Tdap vaccination details to discuss why we choose Tdap patients’ data as the reference system.

**Methods**

This section will introduce the data source used in the study first. Following this, we delineate our study's three key dimensions: Demographic analysis, Temporal analysis, and Spatial analysis. Each dimension offers a distinct analytical perspective. The study was approved by Stony Brook Cancer Center and the dataset was sourced from the administrative database of NYS.

Data Sources

This study utilized patient vaccination data from NYSIIS (New York State Immunization System) which is a dataset managed by the New York State Department of Health, collecting immunization records for individuals from 2008 to prevent vaccine-preventable diseases among children in the state. Our research scale focus on the Long Island (Suffolk County and Nassau County). We filtered data using “Gardasil” to identify patient visits for the HPV dose. Due to experimental requirements, we also gathered data Tdap (Tetanus, Diphtheria, and Pertussis) vaccine relevant data. Only Long Island patients (with a valid Long Island 5 digital zip code) who either received HPV or Tdap were included in our analysis. In total: 1756134 patients’ visit records, and 888277 unique patients were included in our analysis. The dataset also contained patient-level information, which include demographics, vaccination trade, vaccination date, provider information, and patients’ zip code, among other details.

A list of Tdap trade name was used to identify the vaccination records for Tdap patients. Tdap data here serves as the reference system of our HPV vaccination data. Here we use the Tdap data to estimate the total population in each age range on the long island.

[TODO: add list of Tdap trade name on the appendix].

Demographic and Temporal Analysis

All the patient-level data were encrypted. We aggregated the individual data by the client-ID and sorted them by the vaccination data. Only the earliest visits will be counted if a patients have multiple visits. In our study, we utilize patient-level data from 2008 to 2021. Our target group consists of children aged 9 to 13. In order to include all the Tdap data in this timeframe, we need to extend the year range to include 2008 to 2011, as some children may have received their Tdap vaccinations before 2008, and NYSIIS did not collect this information.

We will conduct a comprehensive analysis focusing on examining gender and race disparities among those vaccinated patients. Additionally, we will concern the seasonal patterns of vaccine visits to explore some climate impact on people’s preferences. The Linear Trend analysis will also be added to support the significance of our study.

[TODO Linear Trend analysis]

All the patient-level data were aggregated by the vaccine year and from those time-series data we will visualize the trend and seasonality. The number of kids from 9 to 13 either receive Tdap or receive the HPV in a specific year was counted as the denominator and the nominator was the number of kids from 9-13 who receive the HPV. The result will show the HPV rate on the Long Island from 2012 to 2021.

Spatial Analysis

This study’s scale is on the Long Island, and we visualized the zip code level spatial distribution of HPV rate calculated by the previous parts using the geo-map. We utilized the shade of color in each block to represent the HPV rate.

[TODO Plan to utilize the TIGER/Line to evaluate the correlations between sociodemographic factors with HPV rates at zip-code level]

**Results**

Overall Rate Analysis

Geographic patterns of HPV rate at zip code level

**Discussion**

Limitations of this study

**Funding/Support:**

**Role of the Funder/Sponsor:**

**Dissemination to participants and related patient and public communities:**

**References**

1. Christie C BC, Cooper R, Kennedy PJ, Madras B, Bondi P. The President's Commision on Combating Drug Addiction and the Opioid Crisis. In. <https://www.whitehouse.gov/sites/whitehouse.gov/files/images/Final_Report_Draft_11-1-2017.pdf>: White House; 2017.

**Figure Legends**

Figure 1.

**A graph of vaccinated population

Description automatically generated**

Figure 1. Demographic disparities of HPV rate on the Long Island, New York State, 2012-2021. The rates are calculated based upon per year HPV or Tdap vaccinated patient counts as the denominator and the HPV patients counts as the numerator.

(B) HPV Rate on the Long Island

1. Vaccinated Populations and HPV portion

**A graph of vaccination rate

Description automatically generated**