Supplemental Online Content

Tabibi T, Barnes JM, Shah A, Osazuwa-Peters N, Johnson KJ, Brown DS. Human papillomavirus vaccination and trends in cervical cancer incidence and mortality in the US. *JAMA Pediatr.* Published online November 29, 2021. doi:10.1001/jamapediatrics.2021.4807

eMethods.
eReferences.

This supplemental material has been provided by the authors to give readers additional information about their work.

eMethods

Study sample

The combined Surveillance, Epidemiology, and End Results (SEER) and National Program of Cancer Registries (NPCR) database collect cancer incidence information for all 50 states and Washington D.C. Data were accessed via SEER*Stat software. The National Center for Health Statistics (NCHS) collects cancer mortality data for all recorded cancer deaths in the 50 states and Washington D.C. Cause of death for the purposes of the NCHS is based on ICD codes on death certificates. Data were also accessed via SEER*Stat software.

The rates of receiving ≥ 3 doses of the HPV vaccine in 2011 were 34.8%, 20.5%, 8.7%, and 0.74% in 13-17, 18-24, 25-29, and 30-39 year old females, respectively. The rates for 18-39 year olds are based on the first author's (JMB) analysis of the National Health Interview Survey (NHIS) 2011 data, and the rates in 13-17 year olds were previously published, based on NHIS-Teen data.[1]

Due to data suppression for event counts < 10, analyses adjusted for regional covariates were not feasible.

Estimation of confidence intervals and P-values in primary analyses

The confidence intervals (CIs) and P-values are based on standard errors that were computed using simulation studies. To account for the variance of the rates and correlation between observations over time in the same age group, we conducted simulation studies using 1,000,000 random draws from bivariate normal distributions for age groups $k \in$(15-24, 25-29, 30-39) with vector of means = ($\text{Rate}_{\text{Age group } k, \text{ pre-vaccination period}}$, $\text{Rate}_{\text{Age group } k, \text{ post-vaccination period}}$) and

covariance matrix = S * R * S. R was the correlation matrix and S is the vector of standard deviations = standard errors = ($SE_{\text{Age } group\ k, \text{ pre-vaccination period}}$, $SE_{\text{Age group } k, \text{ post-vaccination period}}$). While the normal distribution may be a good approximation (as the rate can be viewed as a mean, which is asymptomatically normally distributed), there is some degree of right skew, most apparent with small numbers of events / low rates. To avoid using standard error estimates that were too small, we estimated the standard error based on the skewed right side of the distribution, defining SE = (95% CI upper bound – rate)/ 1.95996. Note that there was some concern that the normal distribution may not be appropriate since there is no lower bound of 0; however, the probability of the distribution density being <0 was extremely small (e.g. 0.00063 for the 2016-18 mortality rate for 15-24 year olds, 5e-17 for the 2010-17 mortality rate for 15-24 year olds, and effectively 0 for every other distribution in the above analyses). The correlation ($Rate_{\text{Age group } k, \text{ pre-vaccination}}$, $Rate_{\text{Age group } k, \text{ post-vaccination}}$) was estimated in a simulation study, calculating the correlation between vectors $\mathbf{Rate}_{\mathbf{pre, age\ group\ k}}$ and $\mathbf{Rate}_{\mathbf{post, age\ group\ k}}$, with element $Rate_{\text{pre, age group k,i}}$ and $Rate_{\text{post, age group k,i}}$ equal to two randomly selected single-year rates from age group *k* (with replacement), defining the earlier year as the "pre-" year and the later year as the "post-" year, for $i \in (1,2,\ldots,500)$. After the draws were obtained, the following calculations were performed: $\%\ Change_{\text{Age group } k} = (Rate_{\text{Age group } k, \text{ post-vaccination period}} - Rate_{\text{Age group } k, \text{ pre-vaccination period}})/ Rate_{\text{Age group } k, \text{ pre-vaccination period}}$; $\text{Relative change} = \%\ Change_{\text{Age group 15-24}} - \%\ Change_{\text{Age group } k}$. The distributions of the resultant estimates were utilized to obtain the corresponding confidence intervals (2.5% and 97.5% percentiles of resultant distribution) and P-values (smallest value α such that a $(1-\alpha)*100\%$ confidence interval contains 0).

Testing of the parallel trends assumption

A fundamental assumption of difference-in-differences analyses is the parallel trends assumption. While we do not perform a traditional difference-in-differences analysis since we use % change rather than absolute percentage point changes, we still performed tests to ensure that the percent changes in cervical cancer incidence in the period prior to HPV vaccine introduction were similar between the age groups studied. We gathered yearly incidence and mortality rates. We then compared the year-to-year percent changes in 15-24 and 25-29 or 30-39 year old females from 2001-2005 (e.g. 2001 to 2002, 2002 to 2003, etc.), the pre-vaccine period, as well as from 2001-05 to 2006-09, where we would not expect any differential changes in rates between the age groups. Note that the calculations, including variances, for these year-to-year changes are identical to the primary analyses, except that the "pre" and "post" periods are different. However, note that the formal tests of parallel trends are underpowered since they are based on testing the null hypothesis that the changes in mortality were the same between state groups rather than a test that the differential change was less than a pre-specified margin (see articles by Bilinski and Hatfield, Khan-Lang and Lang, and Roth for additional insights and discussion).[2–4]

Parallel trends testing revealed that the relative difference in the percent change in cervical cancer incidence differed for 15-24 and 25-29 year old females from 2001-2002 (27.44, 95% CI = 1.96 to 57.94, p=.034), but otherwise there were no relative differences in the percent change from 2002-2003 (-7.07, 95% CI = -27.21 to 15.34, p=.52), 2003-2004 (5.27, 95% CI = -16.71 to 30.38, p=.65), 2004-2005 (12.91, 95% CI = -11.38 to 40.91, p=.31), 2001-05 to 2006-09 (-2.0, 95% CI = -10.42 to 6.68, p=.65), and 2004-05 to 2006-09 (-10.83, 95% CI = -22.2 to 1.3, p=.079). Similarly, the relative difference in the percent change in cervical cancer incidence was significantly different for 15-24 and 30-39 year-old females from 2001-2002 (23.94, 95% CI =

0.45 to 53.26, p=.045) and from 2001-05 to 2006-09 (-8.54, 95% CI = -16.26 to -0.36, p=.041), though a revised pre-vaccine time period showed no differences in the percent change (2004-05 to 2006-09; -8.87, 95% CI = -18.34 to 1.98, p=.10). There were no relative differences in the percent change in cervical cancer incidence between 15-24 and 30-39 year-old females from 2002-2003 (-12.4, 95% CI = -29.7 to 8.17, p=.22), 2003-2004 (-2.41, 95% CI = -21.84 to 21.13, p=.83), or 2004-2005 (9.78, 95% CI = -11.71 to 36.08, p=.40). There were no relative differences in the percent change in cervical cancer mortality between 15-24 and 25-29 year-old females from 2001-2002 (46.18, 95% CI = -42.97 to 251.64, p=.35), 2002-2003 (-32.2, 95% CI = -111.55 to 77.3, p=.46), 2003-2004 (1.47, 95% CI = -77.34 to 139.05, p=.97), 2004-2005 (18.79, 95% CI = -29.57 to 98.56, p=.48), or 2001-05 to 2006-09 (-10.2, 95% CI = -42.41 to 26.81, p=.56). There were also no relative differences in the percent change in cervical cancer mortality between 15-24 and 30-39 year-old females from 2001-2002 (40.52, 95% CI = -36.4 to 245.54, p=.37), 2002-2003 (1.01, 95% CI = -50.73 to 105.01, p=.98), 2003-2004 (20.72, 95% CI = -41.52 to 155.8, p=.58), 2004-2005 (-25.07, 95% CI = -69.05 to 53.83, p=.41), or 2001-05 to 2006-09 (-18.99, 95% CI = -45.83 to 15.46, p=.24).

In summary, the parallel trends assumption was met for all incidence and mortality rate analyses with the exception of divergent incidence rate trends from 2001-02. However, this was not considered an issue given similar incidence trends over other years and similar results in a sensitivity analysis with a revised pre-vaccine period (2004-05) using 2010-17 as the post-vaccine period: % change for 15-24 years, -37.58 (95% CI = -43.40 to -30.65, p<.001); % change for 25-29 years, -7.92 (95% CI = -13.44 to -1.74, p=.013; $P_{\text{relative change}}$<.001); % change for 30-39 years, -7.57 (95% CI = -10.23 to -4.77, p<.001); $P_{\text{relative change}}$ <.001). Note that the

parallel trends assumption was satisfied for comparisons of incidence rates from 2004-05 to 2005-09.

eReferences

1. Curtis CR, Dorell C, Yankey D, et al. National Human Papillomavirus Vaccination Coverage Among Adolescents Aged 13–17 Years — National Immunization Survey – Teen, United States, 2011. *Morb Mortal Wkly Rep*. 2014;63(2):61-70. Accessed May 26, 2021. https://www.cdc.gov/mmwr/preview/mmwrhtml/su6302a10.htm
2. Bilinski A, Hatfield LA. *Nothing to See Here? Non-Inferiority Approaches to Parallel Trends and Other Model Assumptions* *.; 2019. Accessed May 12, 2021. https://arxiv.org/pdf/1805.03273.pdf
3. Roth J. *Pre-Test with Caution: Event-Study Estimates After Testing for Parallel Trends*.; 2021. Accessed May 12, 2021. https://jonathandroth.github.io/assets/files/roth_pretrends_testing.pdf
4. Kahn-Lang A, Lang K. The Promise and Pitfalls of Differences-in-Differences: Reflections on 16 and Pregnant and Other Applications. *J Bus Econ Stat*. 2020;38(3):613-620. doi:10.1080/07350015.2018.1546591