Effects of Vaccination Rates on Pertussis Cases in California Counties

Albert Kuo

5/10/2019

1 Abstract

We found that counties in California with higher kindergarten vaccination rates tend to have lower rates of pertussis. With every 1% increase in vaccination rates, the rate of pertussis decreases by 3.1% (95% CI: -0.6%, 6.6%). Our analysis adds to existing studies that support the importance of immunizations as a preventative measure.

2 Introduction

Vaccinations provide immunity to diseases and are one of the most important public health strategies to prevent outbreaks. In 1998, a paper linked the Measles/Mumps/Rubella (MMR) vaccine to autism, instigating concerns about the safety of vaccines and a decline in vaccination rates in the United States. Although the paper has since been retracted, its consequences continue to have a serious impact on public health. Outbreaks of diseases in California in the early 2010s have been linked to reduced vaccination frequency. In this analysis, we investigate whether Diphtheria/Tetanus/Pertussis (DTP) vaccination rates among kindergarteners are associated with rates of pertussis, also known as whooping cough, in California counties.

3 Methods

3.1 Imputation of County Populations

We estimated county populations for every year from 2010 to 2014 by dividing the number of pertussis cases (p) by the rate (r). If the number of pertussis cases is 0 for a given county and year, then their population is imputed by the average of that county's populations in other years. If a county did not have any pertussis cases in any year and we are therefore unable to get a good estimate of their population, then they are not included in our model.

3.2 Poisson Regression

A Poisson generalized linear model was used to analyze the relationship between the DTP vaccination rate and pertussis rate among counties in California. The model takes the following form:

$$\log(E(p|x)) = X\beta + \log(n)$$
$$\iff \log\left(\frac{E\left(p|x\right)}{n}\right) = X\beta$$

For county i and year t, the model can be written as:

$$\log\left(\frac{E\left(p_{it}|x_i\right)}{n_{it}}\right) = \alpha + \beta x_i$$

The outcome variable p is the number of pertussis cases, which we assume follows a Poisson distribution. To account for the varying populations by county and year, we use an offset term $\log(n)$, which can also be viewed as an independent variable with a coefficient value constrained to 1. The offset term implies that in our model, we assume that the log of the mean pertussis rate is linearly related to the independent variables. The independent variable x is the cumulative mean rate of DTP vaccination from 2000 up to and including the year for which the pertussis rate was recorded. The rationale behind this is that the effects of vaccination rates persist over the years; a child who is not vaccinated in kindergarten will be at risk of developing pertussis throughout his lifetime.

The null hypothesis is that there is no relationship betwen the cumulative mean vaccination rate and the pertussis rate: $\beta = 0$. Given a 1% increase in the vaccination rate, the pertussis rate ratio is $\frac{E(p|x+0.01)/n}{E(p|x)/n} = \exp(0.01 * \beta)$ and the percent change in pertussis rate is thus $1 - \exp(0.01 * \beta)$.

3.3 Generalized Linear Mixed Model (GLMM)

To account for observed correlations in pertussis cases by county and year (Figure 1), random intercepts u_t and u_i were added and a generalized linear mixed model was fitted. The correlated data can plausibly be explained by the presence of external environmental factors that change every year and affect the overall incidence of disease, as well as county-specific factors that impact incidence. The model takes the following form:

$$\begin{aligned} &\log(E(p|x)) = X\beta + Zu + \log(n) \\ \iff &\log\left(\frac{E(p|x)}{n}\right) = X\beta + Zu \end{aligned}$$

where X are the fixed effects and Z are the random effects. For county i and year t, the model can be written as:

$$\log\left(\frac{E(p_{it}|x_i)}{n_{it}}\right) = \alpha + \beta x_i + u_t + u_i$$

3.4 Generalized Estimating Equations (GEE)

We also use generalized estimating equations to account for correlated data. As in Poisson regression and GLMM, we use the same link function and assumption of a Poisson distribution on the outcome:

$$\log\left(\frac{E(p|x)}{n}\right) = X\beta$$

Since correlations by year appear to be more predominant than correlations by county (Figure 1), we choose to model the correlation by year using generalized estimating equations. The variance structure was assumed to be unstructured, since it is plausible that the correlations between pairs of counties vary depending on the counties.

4 Results

The dataset includes kindergarten DTP vaccination rates in California counties from the years 2000 to 2015 and the number of pertussis cases from the years 2010 to 2014.

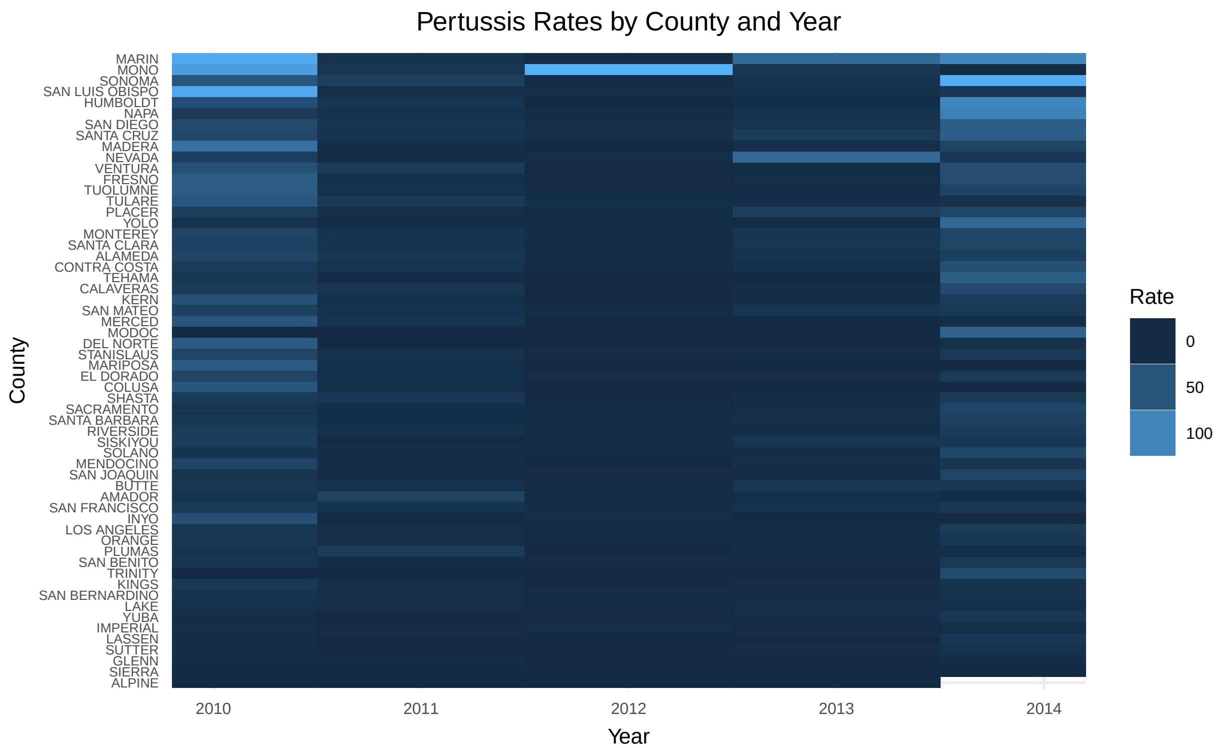

Figure 1: There were higher pertussis rates (per 100,000 people) in 2010 and 2014 in most counties.

4.1 Pertussis Case Rates

In Figure 1, we observe that pertussis case rates are higher in 2010 and 2014, corresponding to known pertussis outbreaks. However, the observed trend varies in both magnitude and pattern by county. Some counties, such as Sonoma and Modoc, experienced high pertussis rates in 2014, but not in 2010. Conversely, other counties, such as San Luis Obispo and Madera, experienced much higher pertussis rates in 2010, but not in 2014. There is also an extreme outlier in Mono, which had an unusually high pertussis rate in 2012, but did not experience the higher rates of pertussis seen in other states in 2014. These results suggest that there may be additional external factors at play that vary by county and subsequently affect their pertussis rates.

4.2 DTP Vaccination Rates

In general, we see that vaccination rates have been steadily declining since 2000 in many counties, with stronger trends evident in counties like Nevada, Tuolumne, and Humboldt (Figure 2). The declining rates in these counties may have contributed to the pertussis outbreaks observed in 2010 and 2014.

4.3 Association between DTP Vaccination Rates and Pertussis Rates

To analyze whether there is a significant association between DTP vaccination rates and pertussis rates, three models were used. The null hypothesis is that there is no relationship between DTP vaccination rates and pertussis rates in California counties. Two counties, Alpine and Sierra, were excluded from the analysis as we were not able to estimate their populations. In Figure 3, we plot the cumulative mean vaccination rate x and the log of the pertussis rate $\log(p/n+1)$. Each point corresponds to a county in a given year (e.g. Riverside in 2011). Under a poisson log-link regression, any relationship between the two variables is assumed to be linear. We observe a slight negative relationship.

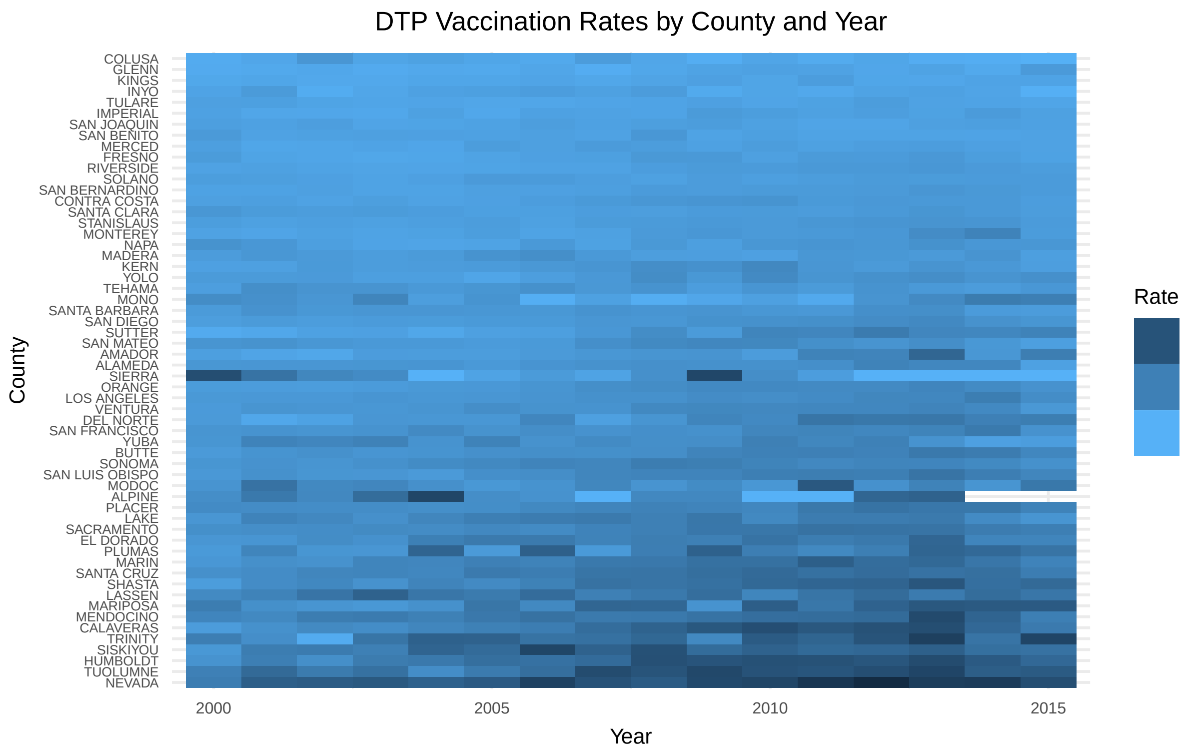

Figure 2: Declining vaccination rates are evident in about half the counties.

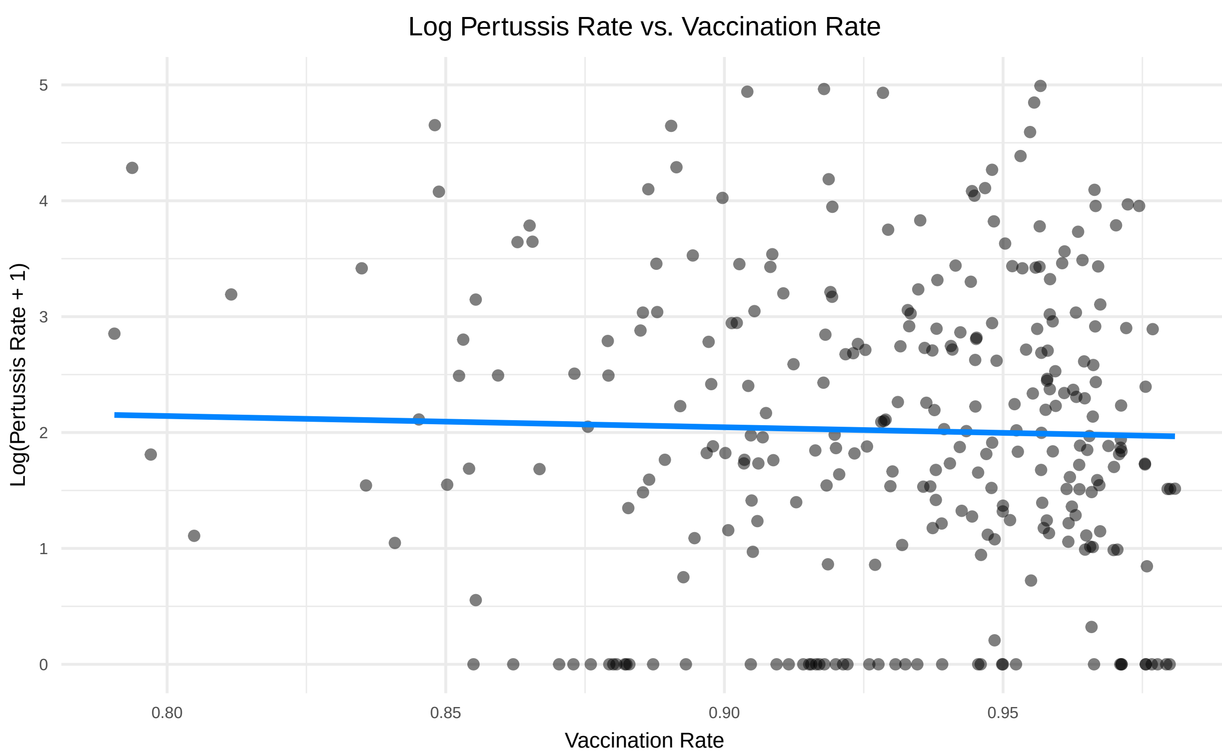

Figure 3: There is a slight negative association between the log pertussis rate and the cumulative mean vaccination rate.

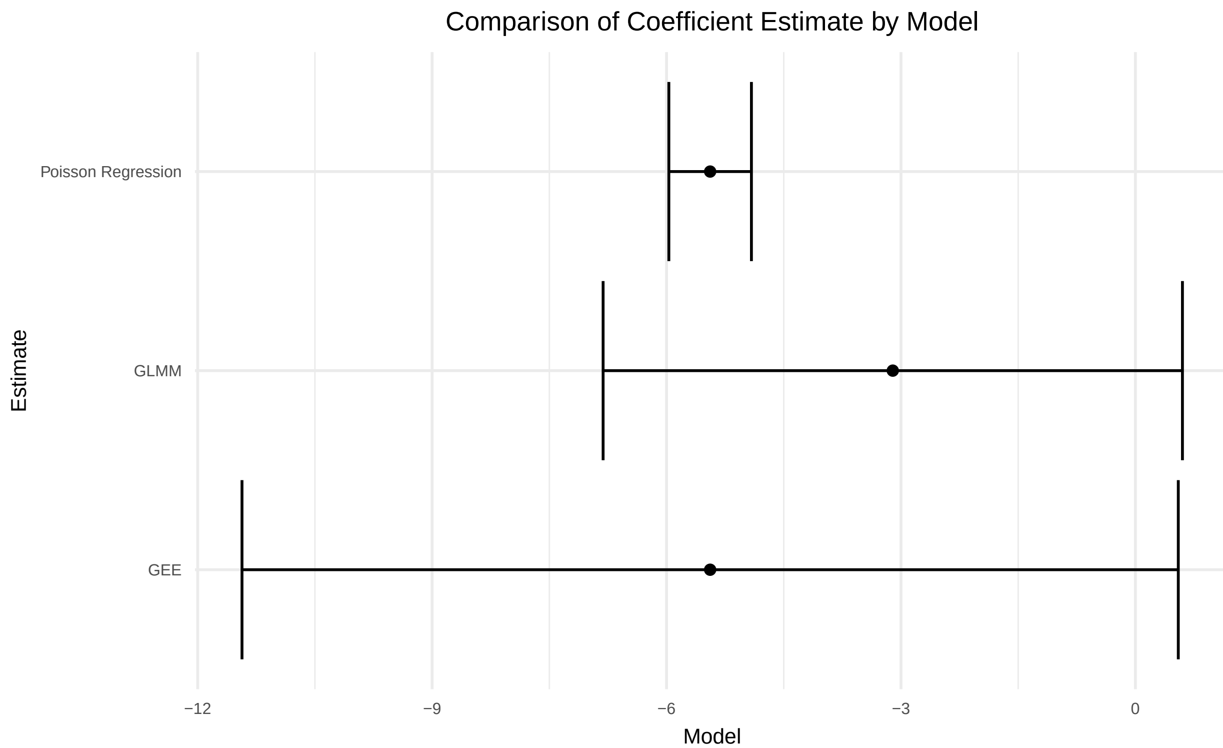

Figure 4: Coefficient estimates and 95% confidence intervals from each model.

While the coefficient for the cumulative mean vaccination rate are comparable in all three models, the GLMM and GEE models estimate a much higher standard error, indicating variance misspecification in Poisson regression (Figure 4). The coefficient for the cumulative mean vaccination rate is estimated to be -3.11 (95% CI: -6.81, 0.60) with the GLMM model and -5.44 (95% CI: -11.43, 0.54) with the GEE model. Since the GLMM model takes into account correlation by both county and year, we believe its results are most reliable. The estimated coefficient under the GLMM model implies that with every 1% increase in the rates of vaccination, there is a decrease of 3.1% (95% CI: -0.6%, 6.6%) in the rate of pertussis. Under a z-test with $\alpha = 0.05$, this association was not found to be significant.

5 Discussion

We found that DTP vaccination rates were negatively associated with pertussis rates in California. A 1% increase in vaccination rates was found to be associated with a 3.1% (95% CI: -0.6%, 6.6%) decrease in pertussis rates. The association is suggestive, but not significant under a p-value of 0.05. The analysis conducted does have a few important limitations. For example, historical data is lacking on pertussis cases prior to 2010 and only five years of pertussis cases were analyzed. Further data collection and follow-up analyses will provide a more robust measure of the relationship between DTP vaccination rates and pertussis cases.