

Yale University

EliScholar – A Digital Platform for Scholarly Publishing at Yale

Public Health Theses

School of Public Health

1-1-2019

Time Series Analysis For Effect Of Vaccines On Pneumonia Or Gastrointestinal Disorders In Brazil

Kang Shen
shenkang212@g.ucla.edu

Follow this and additional works at: <https://elischolar.library.yale.edu/ysphtdl>

Recommended Citation

Shen, Kang, "Time Series Analysis For Effect Of Vaccines On Pneumonia Or Gastrointestinal Disorders In Brazil" (2019). *Public Health Theses*. 1843.
<https://elischolar.library.yale.edu/ysphtdl/1843>

This Open Access Thesis is brought to you for free and open access by the School of Public Health at EliScholar – A Digital Platform for Scholarly Publishing at Yale. It has been accepted for inclusion in Public Health Theses by an authorized administrator of EliScholar – A Digital Platform for Scholarly Publishing at Yale. For more information, please contact elischolar@yale.edu.

**Time series analysis for effect of vaccines on pneumonia or
gastrointestinal disorders in Brazil**

Kang Shen

May 2019

Master of Public Health

School of Public Health

Epidemiology of Microbial Diseases

Committee Chair: Daniel Weinberger, PhD

Committee Members: Virginia Pitzer, ScD

Yale SCHOOL OF PUBLIC HEALTH

Abstract

It is important for public health professionals to understand the effects of vaccine introduction and compare different types of vaccinations. This study will focus on understanding the pre and post vaccination effects of two different infectious diseases: pneumonia and gastrointestinal disorders. One of approaches of getting the accurate estimate of vaccine impact is by comparing rates or trends of the targeted disease in the years after introduction with rates of trend in the year before introduction using a time-series analysis. Specifically, in this study, the synthetic controls method implements an approach to estimating the causal effect of a designed intervention on a time series. With this new approach, potential comparison time series are combined into a composite and are used to generate a counterfactual estimate, which can be compared with the time series of interest, which are effect of vaccines on pneumonia or gastrointestinal disorders after the intervention. After conducting time series analysis on five different regions of Brazil and in national level, this study finds that the introduction of pneumonia vaccine in five regions of Brazil results in a significant increasing in case prevented and decreasing in observed cases compared to synthetic control estimated cases. However, unlike pneumonia vaccine, introduction of gastrointestinal disorder vaccine is not significantly effective for certain region of Brazil in both age groups.

Table of Contents

1. Introduction.....	4
2. Methods.....	5
2.1 Sources of Data.....	5
2.2 Definition of pneumonia hospitalization.....	6
2.3 Definition of intestinal infectious disease hospitalization.....	6
2.4 Study Population.....	6
2.5 Controls.....	7
2.6 Significance.....	8
3. Statistical analysis.....	9
4. Results.....	9
5. Discussion.....	20
5.1 Limitations.....	22
6. Conclusion.....	22
7. EMD Competencies.....	23
8. Appendix.....	24

Introduction

Vaccination, as an epidemiological intervention, is important to keep infectious diseases from spreading. After vaccination, public health practitioners will monitor the infectious disease rate in order to make sure that the vaccine is effective and beneficial to the community. One of approaches of getting the accurate estimate of vaccine impact is by comparing rates or trends of the targeted disease in the years after introduction with rates of trend in the year before introduction using a time-series analysis.^[1]

Before the introduction of vaccination in the United States, rotavirus caused about 2.7 million cases of severe gastroenteritis in children, almost 60,000 hospitalizations, and around 37 deaths per year.^[2] At the same time, also in the United States, pneumonia accounts for 3–18% of all childhood hospital admissions.^[1] For people aged 65 years or older, nearly one million episodes of community-acquired pneumonia occur every year, with 40% resulting in admission.^[3] In the developing countries, pneumonia still remains a leading cause of death across different age groups.^[4] Introduction of a new vaccine is expected to provide direct protection for suspected population by immunizing individuals. The suspected population could also be indirectly protected from herd immunity because of the increasing number of vaccinated individuals. Therefore, the spreading of the infectious disease could be controlled and the admission number of hospitalization caused by infection would be averted significantly. In our study, both effects of rotavirus and pneumococcal vaccines are measured based on the hospitalization data.

Instead of focusing on the seasonal changes of infectious disease, this study is controlled for seasonality while estimating the total effect (direct & indirect) of these vaccines across different age groups based on the data of hospitalization. Data were collected at the national and regional

level of Brazil with the outcomes of either pneumonia or gastrointestinal disorders including both pre and post-vaccination. Comparison of rotavirus vaccine and pneumonia vaccine effects on gastrointestinal disorders and pneumonia will be conducted to generate evidence on public health impact. The aversion of hospitalization admission is used as the end point to estimate the effect of vaccines.

In this study, the synthetic controls method, which was originally developed and named as “Casual impact” in R package, which implements an approach to estimating the causal effect of a designed intervention on a time series. This method is based on the assumption that there is a set of control time series that were themselves not affected by the intervention, which in our study is that control infectious diseases were not affected by the introduced vaccines. With this new approach, potential comparison time series are combined into a composite and are used to generate a counterfactual estimate, which can be compared with the time series of interest, which are effect of vaccines on pneumonia or gastrointestinal disorders after the intervention.^[1]

Methods

Sources of Data

Hospitalization data were obtained from national and regional level data of Brazil across different age groups from the Hospital Information System administrative database of the National Unified Health System (SIH-SUS). We expect to include pre-vaccination and post-vaccination hospitalization data. Targeted individuals were diarrhea-related hospitalization and pneumonia admission, therefore only these two types of data were needed to be collected and analyzed. More specifically, based on the International Classification of Diseases (ICD-10), J12-J18 is the

targeted code for influenza and pneumonia and A00-A09 is the targeted code for intestinal infectious disease from the hospitalization data.

Definition of pneumonia hospitalization

In this study the following ICD-10 codes are used: J12, viral pneumonia; J13, Streptococcus pneumoniae-caused pneumonia; J14, Haemophilus influenzae-caused pneumonia; J15, bacterial pneumonia not otherwise specified; J16, other infectious organism not elsewhere classified; J17, pneumonia in diseases classified elsewhere; and J18, unspecified organism.^[5]

Definition of intestinal infectious disease hospitalization

In this study the following ICD-10 codes are used: A00, Cholera; A01, Typhoid and paratyphoid fevers; A02, other salmonella infections; A03, Shigellosis; A04, Other bacterial intestinal infections; A05, Other bacterial foodborne intoxications, not elsewhere classified; A06, Amebiasis; A07, Other protozoal intestinal diseases; A08, Viral and other specified intestinal infections; A09, Infections gastroenteritis and colitis, unspecified.^[5]

Study Population

The present investigation included children aged under two years old living in Brazil, hospitalized due to all causes from January 2010 to December 2013, in all five different regions of Brazil: North, Northeast, Center-West, Southeast and South.

Controls

Table and definition for controls of synthetic control estimation model

ICD-10 Code for Control	Definition
A17	Tuberculosis of nervous system
A18	Tuberculosis of other organs
A19	Miliary tuberculosis
A39	Meningococcal infection
A41	Other sepsis
B20_24	Human immunodeficiency virus [HIV] disease
B34	Viral infection of unspecified site
B96	Other bacterial agents as the cause of diseases classified elsewhere
B97	Viral agents as the cause of diseases classified elsewhere
B99	Other and unspecified infectious diseases
C00_D48	Neoplasms
D50_89	Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism
E00_99	Endocrine, nutritional and metabolic diseases
I00_99	Diseases of the circulatory system
J20_J22	Other acute lower respiratory infections

K00_99	Diseases of the digestive system
L00_99	Diseases of the skin and subcutaneous tissue
M00_99	Diseases of the musculoskeletal system and connective tissue
N00_99	Diseases of the genitourinary system
P00_99	Certain conditions originating in the perinatal period
Q00_99	Congenital malformations, deformations and chromosomal abnormalities
S00_T99	Injury, poisoning and certain other consequences of external causes
V00_Y99	External causes of morbidity
Z00_99	Factors influencing health status and contact with health services

Significance

Synthetic control model is used to detect and adjust for unmeasured bias and confounding in the evaluation of vaccine programs.^[1] Using data from regional level of Brazil in North, Northeast, Center-West, Southeast and South, the synthetic control model can evaluate the changing of hospitalization due to all-cause pneumonia and gastrointestinal disorders with the introduction of corresponding vaccines. Comparing to other linear regression and time series model, this method is likely to increase the accuracy and effectively adjust for unexplained trends in the

hospitalization data from these five regions of Brazil.

Statistical analysis

Hospitalization data was analyzed through synthetic control method in order to adjust for seasonal changes and unexpected trends. Because many unexpected factors aside from targeted vaccination could influence the disease rate, it is necessary to use control diseases to adjust for unrelated trends. First of all, linear regression model will be used to fit the data from pre-vaccine hospitalization for establishing a relationship between control and targeted diseases. After that, post-vaccine data for control disease will be plugged in to predict the estimated targeted disease value. Finally, the estimated targeted disease hospitalization data will be compared to observed hospitalization data to obtain the difference. Analyzed results will yield hospitalization incidence changing in all states of Brazil in different age groups.

Key assumptions:

1. None of the control diseases are influenced by these two vaccines.
2. The relationship between targeted diseases and the control diseases do not change over time.

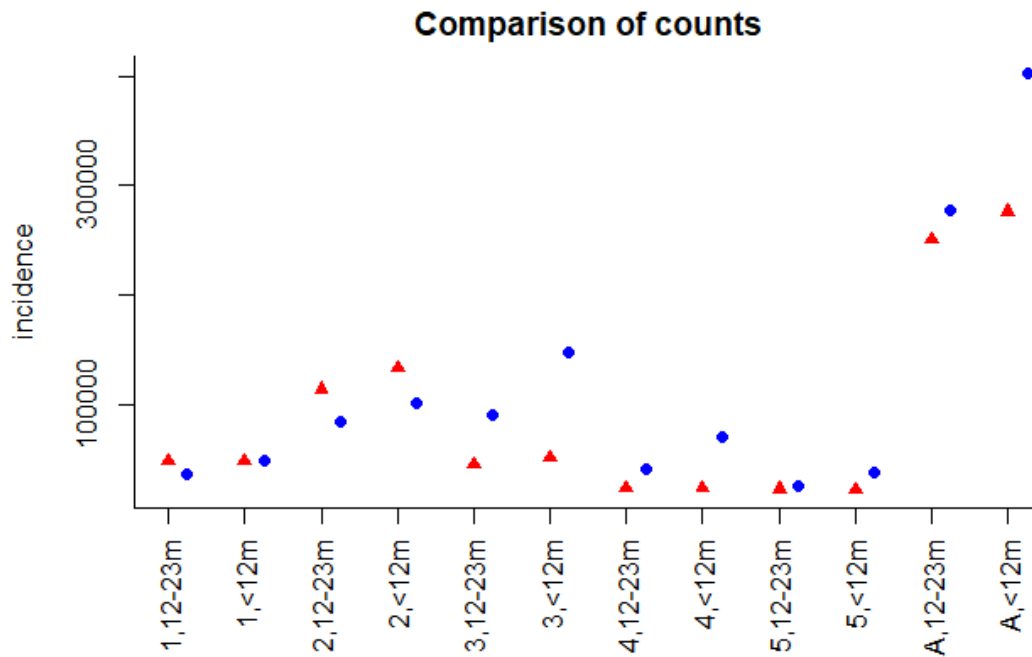
Results

We expected to observe changes in incidence from time series analysis generated results in certain age groups. It is possible that different age groups present different scenarios in the time series. The differences in hospitalization decline after introduction of vaccines between states in

Brazil are likely to be detected from the evidence generated by time series analysis. More importantly, the declination of hospitalization for gastrointestinal disorders and pneumonia could be different due to underlying factors like geographic locations, suspected populations, transmission rate and more.

Comparison of incidence counts in regional and national level of Brazil:

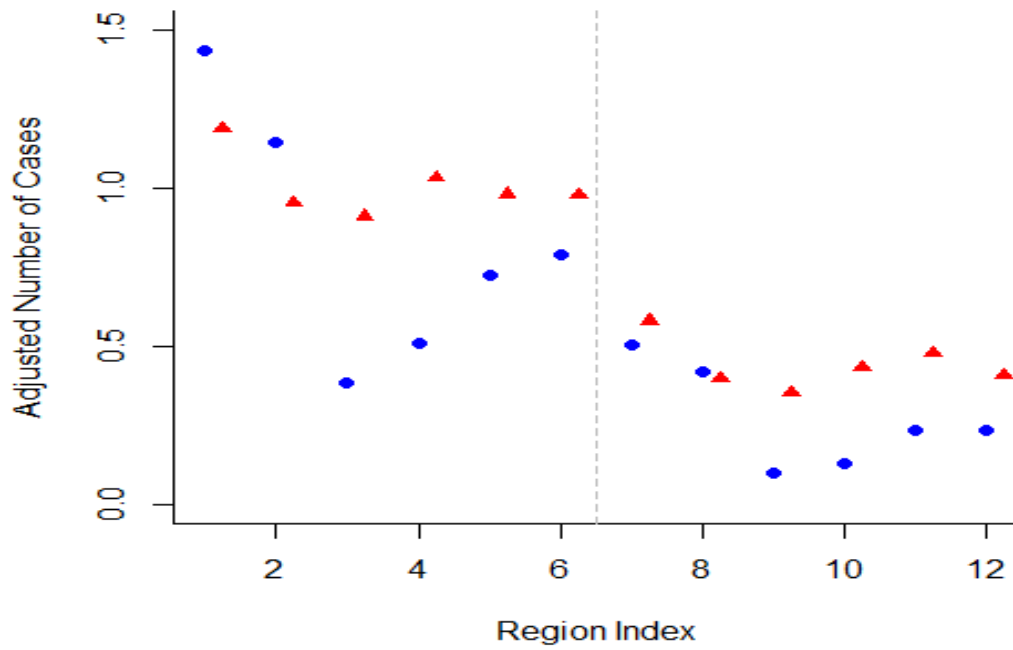
This is the graph to illustrate the different incidence counts of both infectious diseases in different regions of Brazil and in two different age groups under two years old. Specifically, the red triangles represent the incidences of gastrointestinal disorders while the blue dots represent the incidences of pneumonia. The number on the x-axis represents the region number of Brazil, at the same time A represents the national level incidences of both diseases. According to this graph, region one incidence counts of both infectious diseases are close to each other with amount around 50000; region two the incidences of gastrointestinal disorders, which are about 120000, are higher than the incidences of pneumonia, which are about 100000; region three the incidences of pneumonia are higher than the incidences of gastrointestinal disorders in both age groups, with larger difference in less than 12 months age group; region four overall the incidences of pneumonia are higher than the incidences of gastrointestinal disorders in both age groups as well; region five both incidences of two diseases have similar lower counts than other regions of Brazil. In the national level, the incidences of both diseases are higher in less than 12 months age groups. Additionally, incidences of pneumonia are higher than gastrointestinal disorders nationally in Brazil for patients under two years old according to this study.



Comparison of Adjusted number of cases in regional and national level of Brazil:

This graph is adjusted for the overall hospitalization data of Brazil for pneumonia and gastrointestinal disorder incidences. Similarly, the red triangles represent the incidences of gastrointestinal disorders while the blue dots represent the incidences of pneumonia. The dashed gray line separated the overall graph into two parts: the right part represents the adjusted number of cases for 12-23 months age group, the left part represents the adjusted number of cases for under 12 months age group. It is obvious to find out that under 12 months age group has lower adjusted number of cases, which is below 1.0, in all regions of Brazil compared to 12-23 months age group. At the same time, the adjusted number of cases of pneumonia is lower compared to the adjusted number of cases of gastrointestinal disorder in two different age groups. The only exceptions are on the left side of the graph, where region one and region two with 12-23 months age group have higher pneumonia adjusted number of cases compared to gastrointestinal disorder. Therefore, before the introduction of vaccination of both infectious diseases, the incidence of pneumonia should be less than that of gastrointestinal

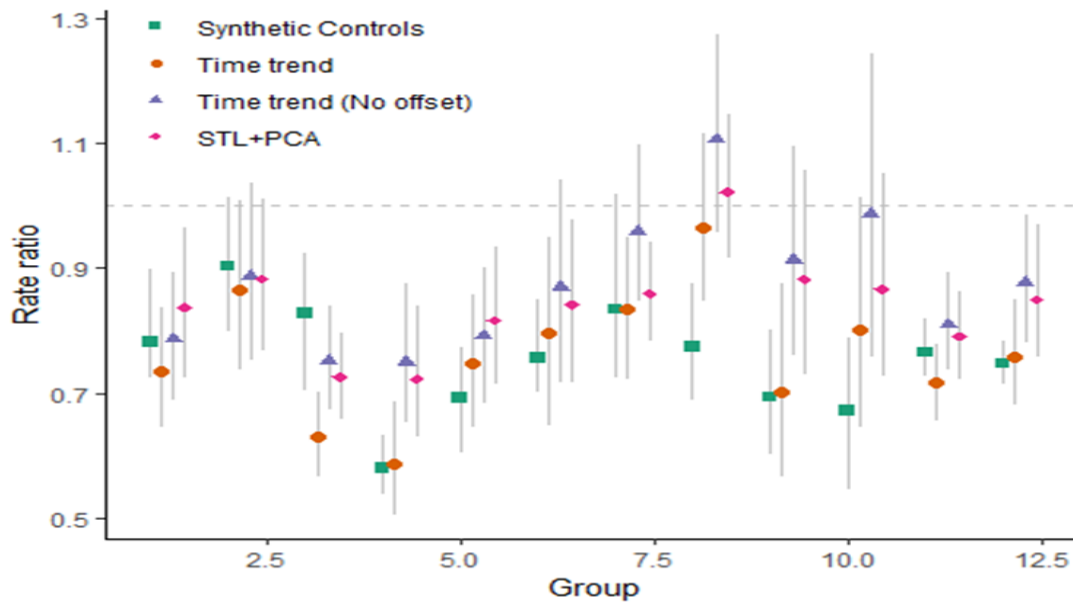
disorders.



Plot of Rate ratios, with size proportional to cross validation weights:

The dash line in this graph is rate ratio equal to one, which means there is no difference of individuals getting infected post-vaccination compared to pre-vaccination. The points below the dash line means that the overall rate ratio is lower than one, which means the number of individuals getting infected post-vaccination is less than number of individuals getting infected pre-vaccination.

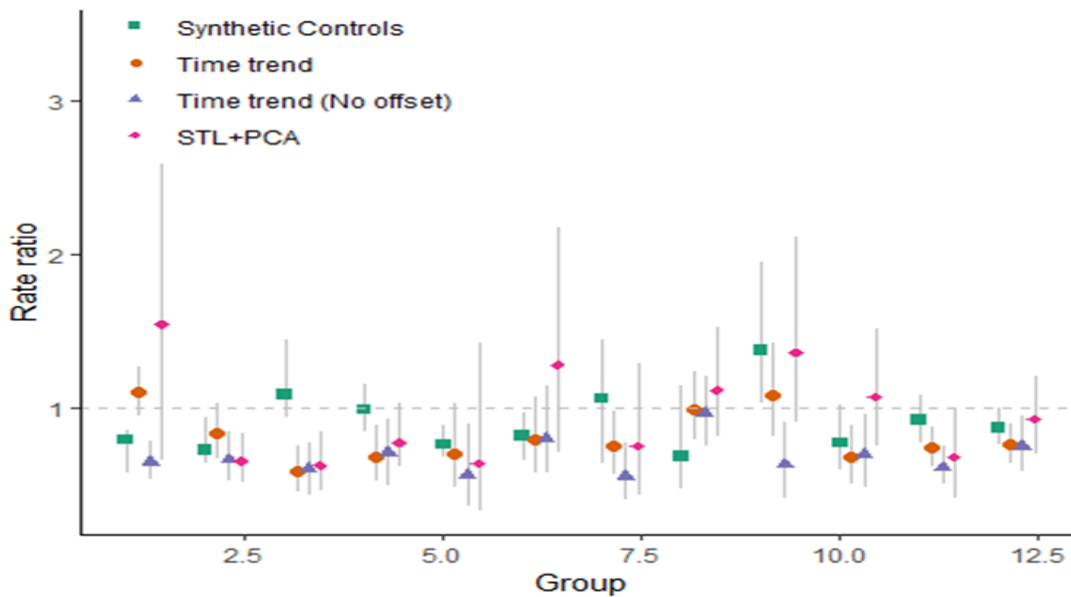
J12_J18:



By checking the synthetic controls model for effect of vaccines on pneumonia, the study shows that the rate ratio for all the age groups in five regions of Brazil is lower than one. At the same time, most of the age groups with four different analysis methods show similar rate ratio lower than one.

A00_A09:

By checking the synthetic controls model for effect of vaccines on gastrointestinal disorders, the study shows that the rate ratios for all the age groups in five regions of Brazil are distributed either above the dash line or below it. Most of the synthetic controls rate ratios are close to one, which is no different between post-vaccination and pre-vaccination. At the same time, most of the age groups with four different analysis methods show similar rate ratio lower than one.



Plot Observed vs expected yearly time series:

In these groups, horizontal black lines represent the expected number of cases derived from synthetic controls estimate, while the black dots are real-life observed number of cases. The gray area contains the black line are the 95% confidence interval for expected number of cases. The vertical dash lines at year 2010 represent the time of introduction of vaccine for different regions of Brazil.

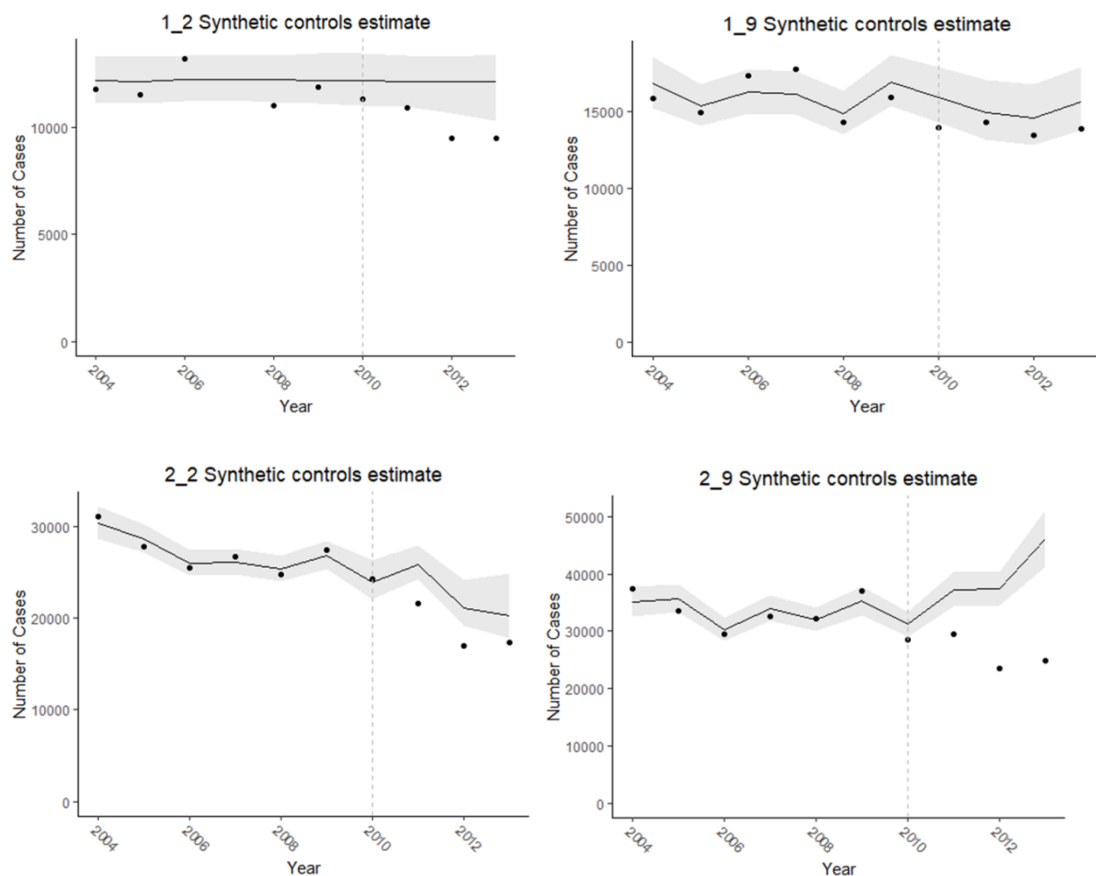
J12_J18:

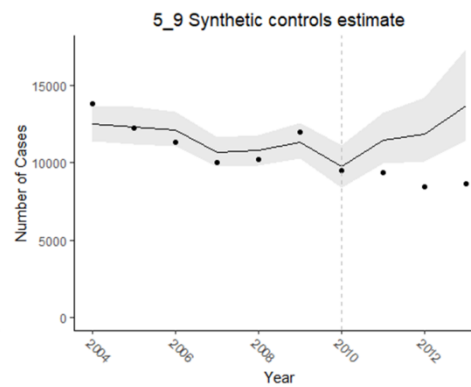
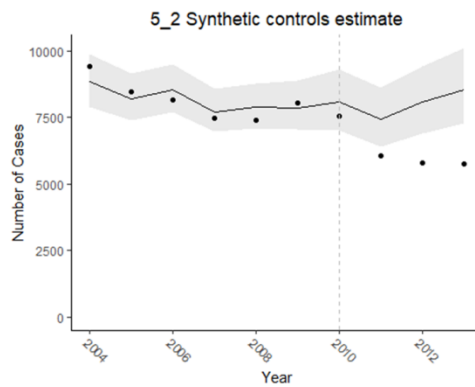
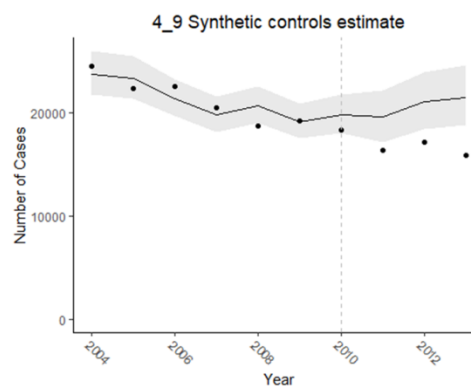
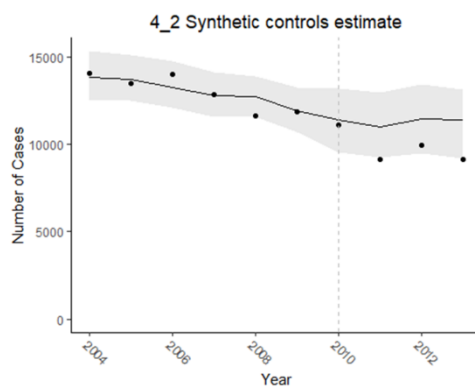
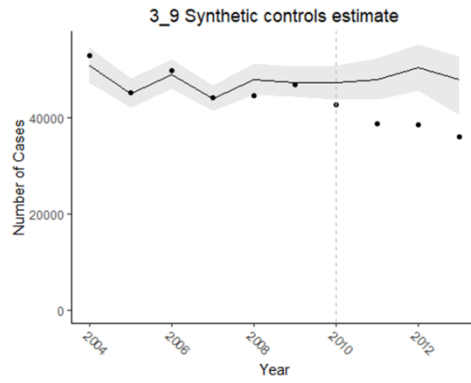
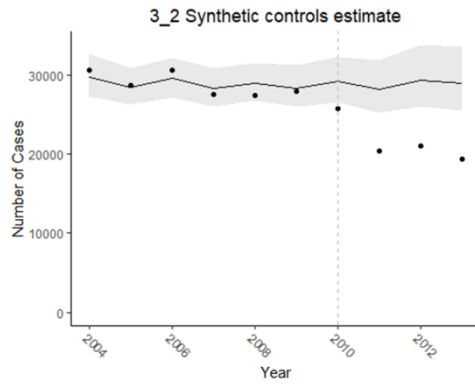
The first region for age group under 2 years old has lower number of cases for observed number of cases compared to expected number of cases derived from synthetic controls estimate after 2010. The observed number of cases for age group 12 month to 23 month is out of the 95% confidence interval of expected number of cases. However, the observed number of cases age group under 12 months is within the 95% confidence interval of expected number of cases.

Similarly for region two, region three and region five, age group under 2 years old has lower

number of cases for observed number of cases compared to the expected number of cases derived from synthetic controls estimate after 2010. Additionally, both observed numbers of cases for age groups are out of the 95% confidence interval of expected number of cases.

At last, region four for age group under 2 years old has lower number of cases for observed number of cases compared to expected number of cases derived from synthetic controls estimate after 2010. The observed number of cases for age under 12 month is out of the 95% confidence interval of expected number of cases. However, the observed number of cases age group between 12 month and 23 month is within the 95% confidence interval of expected number of cases.



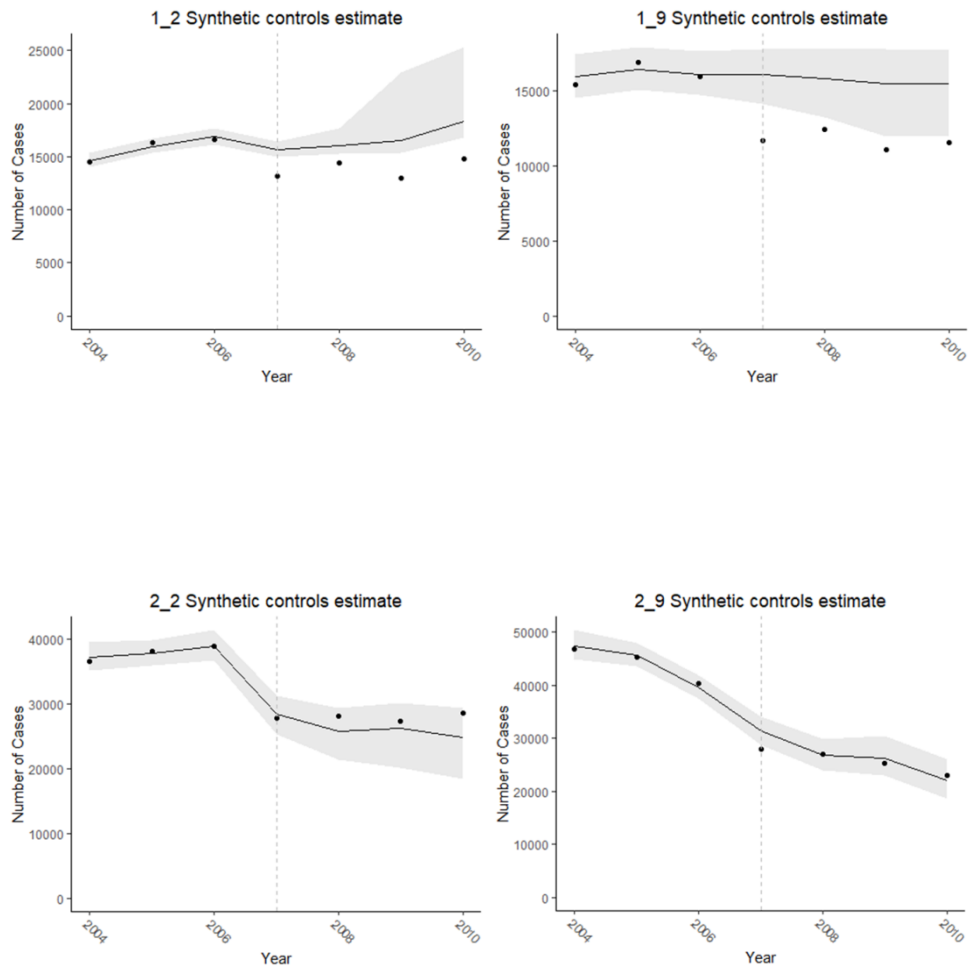


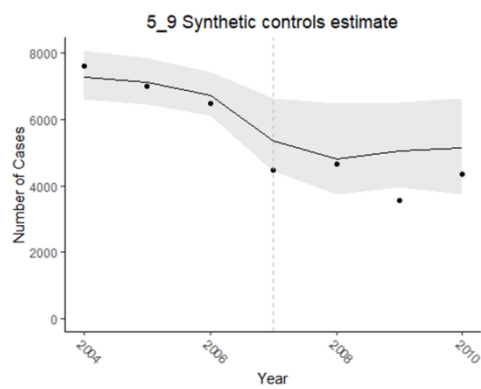
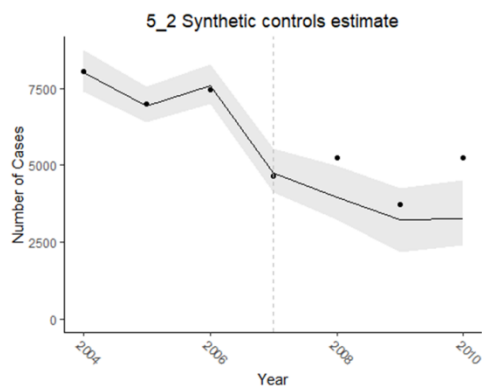
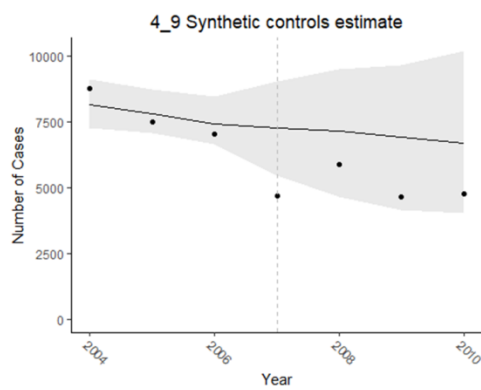
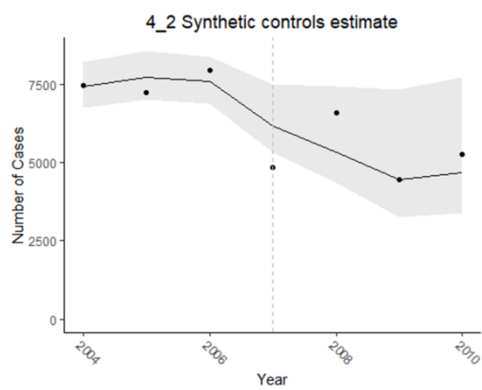
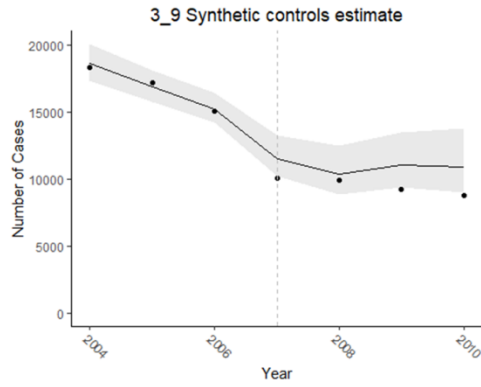
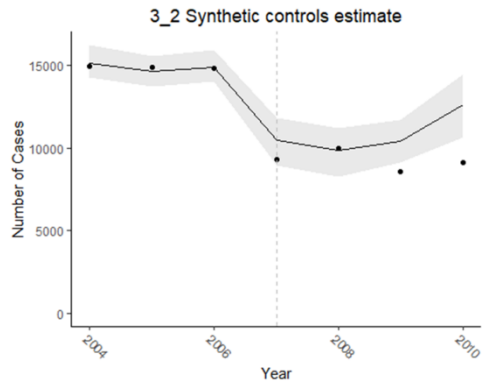
A00_A09:

It is obvious that the confidence interval is much wider after year 2010 for the expected number of cases generated from synthetic control estimated. In general, the first region for age group under 2 years old has lower number of cases for observed number of cases compared to

expected number of cases derived from synthetic controls estimate after 2010. Both of the observed number of cases for age group under 2 years old is out of the 95% confidence interval of expected number of cases.

However, for all the other regions, the differences are not so obvious. For region two, three, four and five, some of the observed numbers of cases are above the expected number of cases generated from synthetic control estimated. Most of the observed numbers of cases are in the 95% confidence interval of the generated expected number of cases.



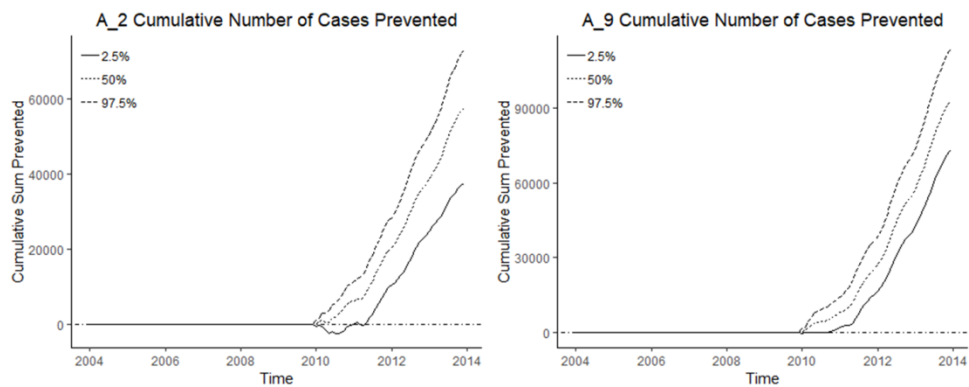


National level cumulative cases prevented for both vaccines:

J12_J18:

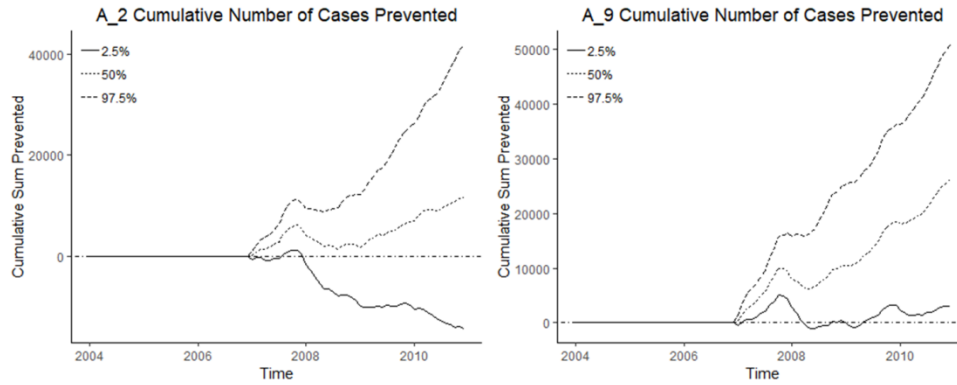
From these two graphs, the cumulative sum of incidence prevented for pneumonia increases

significantly after the introduction of vaccine after 2010, which approximately reaching 60000 for age group between 12 month and 23 month, and reaching over 90000 for age group below 12 month. The 95% confidence interval for both age groups are above the line of zero cumulative sums prevented. At the same time, the 95% confidence interval for the cumulative sum is wider for age group between 12 month and 23 month.



A00_A09:

From these two graphs, the cumulative sum of incidence prevented for gastrointestinal disorders is not increased significantly after the introduction of vaccine after 2010 for age group between 12 month and 23 month because the 95% confidence interval includes the zero cumulative sums prevented in this age group. While for the age group under 12 month, the cumulative sum of incidence prevented increased significantly with 95% confidence interval above the zero line for cumulative sum. At the same time, the 95% confidence interval for the cumulative sum is wider for age group between 12 month and 23 month.



Discussion

Overall, the incidence counts of pneumonia are higher than the incidence counts of gastrointestinal disorders in national and regional level. However, the regional differences were observed as well. The only exception for the incidence counts of gastrointestinal disorders higher than the incidence counts of pneumonia is region two. The incidence counts of pneumonia are also higher in age group less than 12 months compared to age group between 12 month and 23 month, while the incidence counts of gastrointestinal disorders are similar for both age groups in each region and in national level. In general, the prevalence of pneumonia is higher than gastrointestinal disorders in younger patient population according to the hospitalization data. Based on this knowledge, public health professionals and hospitals should prepare more vaccines for pneumonia then vaccines for gastrointestinal disorders for younger patients in Brazil.

At the same time, there is no significant different between post-vaccination and pre-vaccination for gastrointestinal disorders, which there is obvious lower rate ratio difference between post-vaccination and pre-vaccination for pneumonia. The study shows that there is a significant difference between the expected number of cases generated from synthetic control estimated and the observed number of cases after the introduction of vaccines of pneumonia in all five

different regions in Brazil. At the same time, the observed number of cases is generally decreasing each year after the introduction of vaccine in 2010. Between two different age groups, both show similar decreasing patterns after the introduction of vaccines. Unlike vaccines of pneumonia, the introduction of gastrointestinal disorders vaccines effect are only obvious in region one for both age group. There is not a significant difference between the expected number of cases generated from synthetic control estimated and the observed number of cases for region two, three, four and five. Age group of between 12 month and 23 month even shows an increased number of observed cases compared to expected number of cases generated from the analysis. At the same time, the confidence interval after the introduction of vaccines in 2010 is wider for gastrointestinal disorders according to the analysis.

There is a significant increasing of the case prevented after the introduction of vaccine for pneumonia for both age groups, while the cases prevented increase only shows in the age group under 12 months after the introduction of vaccine for gastrointestinal disorders. Therefore, it is obvious that the vaccination process greatly helps the prevention of pneumonia infections for patients under two years old, and the effect of this vaccine is keeping increasing by observing the steady increasing cases prevented over these years. In general, the vaccine for pneumonia is significantly effective and could keep benefiting this population in the Brazil for the upcoming future. However, unlike vaccine for pneumonia, the vaccine for gastrointestinal disorders shows significant differences for two different age groups. It is hard to determine whether the vaccine is effective for age group 12-23 months and under 12 months by considering confidence interval. It is likely that the vaccination is more effective among older patients because patients under 12 months don't expose to too many types of foods, which could potentially cause the

gastrointestinal disorders.

Limitations:

This study only concerns the younger age group, which are under 2-year-old patients with hospitalization data of pneumonia and gastrointestinal disorders. It is possible that other age groups with older patients could show different synthetic control results in different regions of Brazil. Individuals with mature immune system could become less susceptible to infectious diseases like pneumonia and gastrointestinal disorders. Because this study is conducted in a regional and national level instead of state specific level, the geography characteristic is overall more generalized and over-simplified. Geographic difference is an important factor that influencing the prevalence and spreading of infectious diseases. Conditions like number of hospitals, environment hygiene, social economic status, public health facilities and etc. could contribute and impact the effect after the introduction of vaccine in different regions of Brazil. In general, geographic differentiation is too general for this study.

Conclusion

The introduction of pneumonia vaccine in five regions of Brazil results in a significant increasing in case prevented and decreasing in observed cases compared to synthetic control estimated cases. However, unlike pneumonia vaccine, introduction of gastrointestinal disorder vaccine is not significantly effective for certain region of Brazil in both age groups.

EMD Competencies

- Define the scope and worldwide impact of infectious diseases.
- Describe the processes that drive transmission and maintenance of infectious agents.
- Describe the epidemiology of the major infectious diseases worldwide as well as risk exposures and behaviors as these relate to transmission.
- Explain the interrelationship between the environment and the emergence and maintenance of infectious diseases in populations.
- Describe and critically evaluate approaches for the prevention and control of infectious diseases and define the key issues to their effective use.
- Apply principles and concepts obtained through coursework to design and implement studies on the etiology, detection, prevention or control of infectious diseases in the laboratory and field.

Appendix

1. Bruhn, Christian AW, et al. "Estimating the population-level impact of vaccines using synthetic controls." *Proceedings of the National Academy of Sciences* 114.7 (2017): 1524-1529.
2. Farha T, Thomson AH. The burden of pneumonia in children in the developed world. *Paediatr Respir Rev* 2005; 6: 76–82.
3. Fischer TK, Viboud C, Parashar U, Malek M, Steiner C, Glass R, Simonsen L (2007). "Hospitalizations and deaths from diarrhea and rotavirus among children <5 years of age in the United States, 1993–2003". *Journal of Infectious Diseases*. **195** (8): 1117–25. doi:[10.1086/512863](https://doi.org/10.1086/512863). PMID [17357047](https://pubmed.ncbi.nlm.nih.gov/17357047/).
4. Jackson ML, Neuzil KM, Thompson WW, et al. The burden of community-acquired pneumonia in seniors: results of a population-based study. *Clin Infect Dis* 2004; 39: 1642–50.
5. Ruuskanen O, Lahti E, Jennings LC, Murdoch DR (April 2011). "Viral pneumonia". *Lancet*. 377 (9773): 1264–75. doi:[10.1016/S0140-6736\(10\)61459-6](https://doi.org/10.1016/S0140-6736(10)61459-6). PMID 21435708.
6. World Health Organization. International Statistical Classification of Diseases and Related Health Problems, 10th Revision; 2010 (http://www.who.int/classifications/icd/ICD10Volume2_en_2010.pdf). Accessed 15 February 2013.