

## Short Communication

# Bronchiolitis Hospitalization in Southern Brazil from 2002 to 2012: A Count Time-series Approach

**Guilherme Zubatch da Cunha<sup>(1)</sup>, Eniuce Menezes de Souza<sup>(2)</sup>  
and Márcia Lorena Alves dos Santos<sup>(3)</sup>**

[1]. Programa de Pós-Graduação em Medicina Social, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, SP, Brasil.

[2]. Departamento de Estatística, Universidade Estadual de Maringá, Maringá, PR, Brasil.

[3]. Programa de Pós-Graduação em Bioestatística, Departamento de Estatística, Universidade Estadual de Maringá, Maringá, PR, Brasil.

### Abstract

**Introduction:** Because the antibody neutralizing respiratory syncytial virus (anti-RSV) has a short period of immunization and high cost, the identification of regions and months of highest occurrence of bronchiolitis is very important. **Methods:** An Autoregressive Conditional Poisson model was constructed for count data and compared to the standard time-series Poisson regression model. **Results:** The metropolitan area of Paraná presented the highest average occurrence from May to July. **Conclusions:** The constructed model presented a better fit and allowed prediction of when and where the bronchiolitis hospitalizations are distributed.

**Keywords:** Autoregressive Conditional Poisson. Bronchiolitis Hospitalizations. Count Time Series.

Brazil is experiencing an accelerated process of social and demographic changes. Access to health care has been achieved through national health programs and with this, life expectancy at birth has risen from a mere 50 years in the 60s to over 70 in 2020<sup>1</sup>.

However, there are several diseases that create cause for concern, especially those affecting young children that are of great importance for epidemiological research and need to be monitored. Respiratory diseases, one of the main causes of infant mortality around the world, cause 4.5 million deaths per year. Specifically, bronchiolitis is the most common cause of respiratory infections in early childhood and is mainly caused by the respiratory syncytial virus (RSV). Infections caused by RSV occur worldwide and, according to the World Health Organization (WHO), account for about 60 million infections with 160,000 annual deaths worldwide. RSV is responsible for more than 80% of lower respiratory tract infections in children younger than 1 year and is by far the most frequent cause of pediatric bronchiolitis and pneumonia<sup>2,3,4</sup>.

There is no specific treatment for RSV, and some populations of children (newborn, those with some form of congenital heart disease, chronic lung disease, immunocompromized, undernourished, etc.) are at increased risk of morbidity and mortality. The most effective measure is the administration of an antibody, anti-RSV (Palivizumab), which has neutralizing and inhibitory activity against RSV for a period of 30 days. Up to five annual doses of medication, to be administered monthly, are suggested. However, it is recommended that the medication is started one month before the seasonal peak, which is different depending on the region of Brazil. Several factors are responsible for the occurrence of RSV infections, such as geographic location, latitude, and altitude, as well as climatic factors such as temperature, barometric pressure, relative humidity, vapor tension, hours of daylight, precipitation, and dew point<sup>5</sup>. Given that Palivizumab has a high cost to the government, about R\$5,000.00 per bottle, and due to its short period of immunization, it is extremely important to model and identify the months in which the disease's occurrence is more frequent (seasonal peaks).

A criterion for the early detection of bronchiolitis epidemics was developed with the aim of issuing public health alerts and

**Corresponding author:** Msc. Márcia Lorena Alves dos Santos.

**e-mail:** alves.mlorena@gmail.com / mlasantos2@uem.br

**Orcid:** 0000-0002-1098-1944

**Received** 8 August 2018

**Accepted** 13 October 2018

enabling the early implementation of prevention and control strategies<sup>6</sup>. The proposal consists in estimating the basal level of cases of bronchiolitis and predicting using a selected model.

Some investigations of bronchiolitis hospitalization rates have been realized from 2008 to 2010<sup>5</sup>. They have worked with the rate of the number of hospitalizations in relation to the number of live births for each region, which is a common strategy in the literature. However, data on live births need to be available to compute this rate. As the delay in the availability of such data is much longer than that of the number of hospitalizations due to bronchiolitis, it is important to analyze the count data to obtain more up-to-date results.

The aim of this research is to evaluate the number of hospitalizations due to bronchiolitis in health centers in Paraná state from temporal and spatial viewpoints. To account for temporal variations such as trends and seasonal behaviors, as well as other serial correlations, appropriate time-series models for count data were built. Poisson regression models for time series can, in many cases, succeed in modeling this kind of data. However, these models are limited because they assume that events are independent, and the use of these models is still recurrent in the literature<sup>6</sup>. When a count dataset exhibits time dependence, the plain Poisson regression is not adequate<sup>7</sup>. Another model that has been well used in the literature is the Autoregressive Conditional Poisson (ACP) for cases of count data exhibiting autoregressive behavior<sup>8</sup>. An important factor in the decision to use these models is that they are flexible regarding the inclusion of explanatory variables. We aim to show two classes of models for bronchiolitis hospitalization data and compare their performances.

An ecological study of monthly hospitalizations due to bronchiolitis in children younger than 1 year was conducted from 2002 to 2012 in Paraná State, Southern Brazil. The state has a humid subtropical climate in the Northeast and coastal plains and a subtropical climate in South. In 2010, the population of Paraná State was 10,512,349, of which 714,062 (6.9%) were younger than 1 year old. The State is administratively divided into 399 municipalities that are grouped into 22 regional health districts (Instituto Brasileiro de Geografia e Estatística. demographic sense. <http://www.ibge.gov.br>): 1 Paranaíba; 2 Metropolitana; 3 Ponta Grossa; 4 Irati; 5 Guarapuava; 6 União da Vitória; 7 Pato Branco; 8 Francisco Beltrão; 9 Foz do Iguaçu; 10 Cascavel; 11 Campo Mourão; 12 Umuarama; 13 Cianorte; 14 Paranavaí; 15 Maringá; 16 Apucarana; 17 Londrina; 18 Cornélio Procopio; 19 Jacarezinho; 20 Toledo; 21 Telêmaco Borba; 22 Ivaiporã.

Thus, time-series data of the monthly number of patients hospitalized for bronchiolitis for each one of the 22 health districts were obtained from the System of Hospital Information of SUS (SIH-SUS) in the Brazilian Unified Health System database (DATASUS - [www.datasus.gov.br](http://www.datasus.gov.br)) using the 10th revision of the International Classification of Diseases (ICD-10) with the code J21.

The development of the study occurred as recommended by Resolution n. 196/96 of the Brazilian National Health Council. The project was approved by the Ethics Committee of the State

University of Maringá (Legal Report 140/2009) and the Term of Free and Informed Consent was not used because the data were secondary.

Considering that bronchiolitis count data,  $y_t$  follow Poisson distribution and present seasonal patterns, Poisson regression and ACP models were built to take the seasonality into account. This pattern can be dealt with by inserting artificial variables into the construction of the models. Thus, the annual and possible semi-annual (6 months) behaviors were represented by the functions  $\cos(2\pi w_j t/12)$  and  $\sin(2\pi w_j t/12)$  with frequencies,  $w_j = j/12$ ,  $j=1$  and  $j=2$ , respectively, in both the Poisson regression and ACP models for all 22 health districts. The expression of the Poisson regression model fitted to the data is given by,

$$y_t = \beta_0 + \beta_1 t + \gamma_1 \cos\left(\frac{2\pi t}{12}\right) + \gamma_2 \sin\left(\frac{2\pi t}{12}\right) + \gamma_3 \cos\left(\frac{4\pi t}{12}\right) + \gamma_4 \sin\left(\frac{4\pi t}{12}\right) + \varepsilon_t,$$

Where  $\beta_0$  and  $\beta_1$  represent the constant term and the coefficient of the trend, respectively. In addition,  $\gamma_i$ ,  $i=1, \dots, 4$  are the parameters related to the seasonal pattern and  $\varepsilon_t$  represents the random error. For the Conditional Poisson Autoregressive (ACP) model, we have the following expression,

$$y_t = \beta_0 + \beta_1 t + \sum_{j=1}^p \alpha_j y_{t-j} + \sum_{j=1}^q \theta_j \lambda_{t-j} + \gamma_1 \cos\left(\frac{2\pi t}{12}\right) + \gamma_2 \sin\left(\frac{2\pi t}{12}\right) + \gamma_3 \cos\left(\frac{4\pi t}{12}\right) + \gamma_4 \sin\left(\frac{4\pi t}{12}\right) + \varepsilon_t,$$

where  $\alpha$  and  $\theta$  represent the autoregressive and moving-average terms, respectively. Furthermore  $p$  describes the number of lags on the observed variable that are incorporated into the model and  $q$  indicates the lags of previous means.

A group of ten scores was used in this study and each one indicates which model is better adjusted to represent the variability of the data. Two of them are classical: mean absolute error and root mean squared error. The other eight are related to score functions for count model evaluation<sup>9,10</sup>: logarithmic, quadratic, spherical, ranked probability, Dawid-Sebastiani, squared error, mean absolute error, and root squared error scores.

As each score has specific features, we observed that the frequency of the ten scores was in favor of one model in relation to the other for each health division. Due to the spatial variability, maps of the estimated cases from the temporal model were built to improve the visualization of critical regions.

From January 2002 to December 2012, more than ten thousand cases of bronchiolitis were recorded in Paraná State. The 2<sup>nd</sup> health district was the data stream with by far the largest number of cases during the study, with a monthly average of 18.9 hospitalizations and the 22<sup>nd</sup> health district had the fewest, with an average of 0.35 hospitalizations per month. The largest number of cases was reported in May; 96 in the 2<sup>nd</sup> health division, the average cases per month is 3.05, and the median is 1 case per month for all health districts.

All 22 health districts reflected the typical seasonal patterns. From the descriptive analysis, we can see that, in general, more cases occurred between April and July. However, some of the health districts have peaks in other months, thus reinforcing

the need to build a time-series model suitable for each health division.

**Figure 1** shows the adjusted Poisson regression and ACP models for 22 health districts. The trend is not apparent in all districts, being visually noticeable only for some of them, such as for the 2<sup>nd</sup> health district (Metropolitana), where, according to the ACP model, nine and 13 cases were estimated in January and June of 2000, respectively, while these estimates were 14 and 61 in 2012, respectively.

From **Figure 1**, we can see that the ACP model fits the variability of the time series better than the Poisson regression model. However, this was not the case for Poisson models. For most time series, the errors remained autocorrelated. For some health districts, given the nature of the series, there is a certain difficulty in achieving more consistent models such as the 21<sup>st</sup> and 22<sup>nd</sup> health districts.

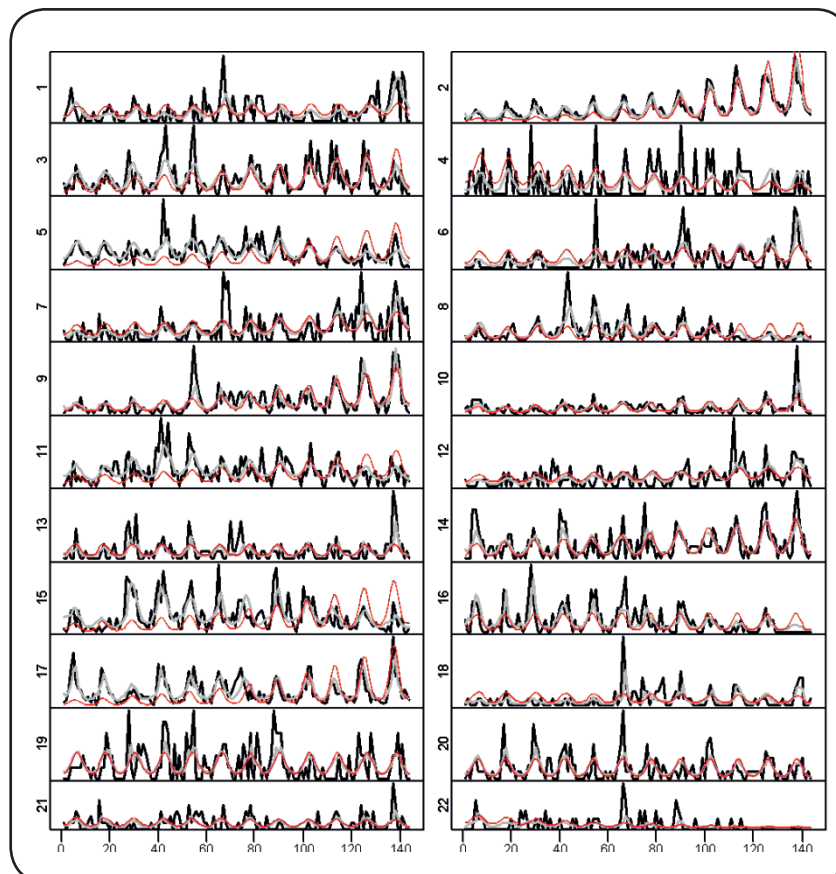
However, in some time series, there is no graphic evidence that one model is better than the other. For this reason, the scores cited in the methodology can help in the comparison of the two models. We computed the frequency of how many of the ten evaluated scores were in favor (superior/equivalent) of the ACP model. The results were very good because 93% of the scores are in favor of ACP model considering all regional health districts. In the few instances the Poisson model was superior to ACP, up

to two of the 10 scores indicated that for nine regional health districts and, only for the Paranaíba health district, three scores were in favor of the Poisson model. Considering the well-known classical RSME score, the ACP model was superior to Poisson model in 100% of the instances.

In **Table 1**, the monthly averages of the estimated values are presented in descending order according to the number maximum of estimated hospitalizations per month.

According to **Table 1**, it is evident that the metropolitan area has far more cases than other regions because of the high concentration of the population. The highest average was found from May to July. In general, a disadvantage in working with the number of hospitalizations rather than rates is that it is not possible to compare these numbers in an absolute form. On the other hand, estimating the number of hospitalizations makes the interpretations direct and is relevant information for a decision regarding which month the public health service should start administering the medicine.

To improve the evaluation and visualization of the spatial epidemiology of bronchiolitis hospitalizations in Paraná state, **Figure 2** presents the spatial/geographical distribution of bronchiolitis hospitalizations estimated using the ACP model. When we present the results using tables and figures we lose some subtle patterns; maps have long been used to describe



**FIGURE 1:** Poisson (red) and ACP (gray) models adjusted to time series (black) of the observed number of bronchiolitis cases.

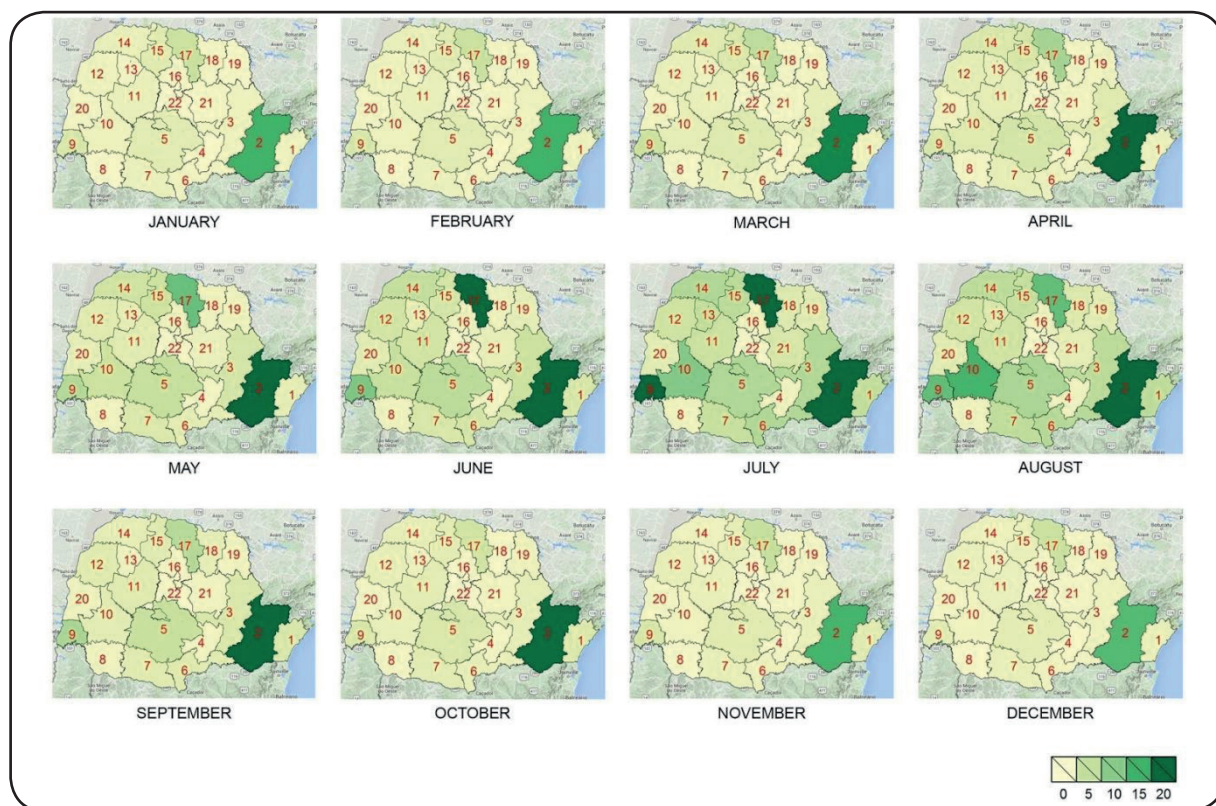


FIGURE 2: Distribution of bronchiolitis hospitalizations estimated from the ACP model for each month of 2012 in Parana State.

TABLE 1: Monthly average of the estimated values from ACP models for the 22 health divisions in Paraná State from 2002 to 2012.

Regional Division	Jan.	Feb.	Mar.	Apr.	May.	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
02 Metropolitana	8	9	15	22	35	<b>40</b>	32	24	15	10	8	7
17 Londrina	5	6	8	12	<b>17</b>	<b>17</b>	12	8	5	4	4	4
05 Guarapuava	3	4	5	7	9	<b>10</b>	<b>10</b>	8	6	5	4	3
15 Maringá	3	4	5	7	<b>9</b>	<b>9</b>	7	5	3	3	2	3
09 Foz do Iguaçu	2	2	3	4	6	<b>8</b>	6	6	3	2	2	2
03 Ponta Grossa	1	2	2	3	5	<b>6</b>	<b>6</b>	5	3	2	1	1
10 Cascavel	1	1	2	3	4	<b>5</b>	<b>5</b>	3	1	1	1	1
11 Campo Mourão	2	2	3	3	4	<b>5</b>	<b>5</b>	4	3	2	2	2
08 Francisco Beltrão	1	1	1	2	3	3	4	4	2	1	1	1
14 Paranavaí	1	2	2	<b>4</b>	<b>4</b>	<b>4</b>	3	2	1	1	1	1
16 Apucarana	1	1	2	2	3	<b>4</b>	2	2	1	1	1	1
20 Toledo	1	1	1	2	<b>3</b>	<b>3</b>	<b>3</b>	2	1	1	0	0
01 Paranaguá	1	1	1	1	1	<b>2</b>	<b>2</b>	<b>2</b>	1	1	1	1
06 União da Vitória	0	0	1	1	<b>2</b>	<b>2</b>	<b>2</b>	<b>2</b>	1	1	1	0
07 Pato Branco	1	1	1	1	<b>2</b>	<b>2</b>	<b>2</b>	<b>2</b>	1	1	1	1
12 Umuarama	1	1	1	1	<b>2</b>	<b>2</b>	<b>2</b>	<b>2</b>	1	1	1	1
13 Cianorte	1	1	1	<b>2</b>	<b>2</b>	<b>2</b>	<b>2</b>	1	1	1	0	0
18 Cornélio Procópio	0	0	1	1	1	1	2	1	1	0	0	0
19 Jacarezinho	0	1	1	1	2	<b>2</b>	<b>2</b>	<b>2</b>	1	1	1	1
21 Telêmaco Borba	0	1	1	1	2	<b>2</b>	<b>2</b>	1	1	1	0	0
04 Irati	0	0	0	0	1	1	1	1	1	0	0	0
22 Ivaiporã	0	0	0	0	1	1	1	0	0	0	0	0

geographic patterns of disease and to transmit visual information regarding the progress of the disease over time and space. In this study, the data were standardized using the maximum of all the series so as to ensure better visualization.

As shown in **Figure 2**, bronchiolitis hospitalizations predicted by the ACP model are from January to December 2012. The maps provide a succinct summary of geographic patterns of the disease. There is no cluster evidence in these maps, and areas of high population tend to be emphasized. The results presented in this research is consistent with the literature in the sense that, in the majority of cases, bronchiolitis is usually seasonal, with epidemics occurring every year<sup>11,12</sup>. We can observe a clear increase in hospitalizations from July until October, where the seasonal peaks occur. On the other hand, the number of cases drops from December to February.

Temporal and spatial variations of diseases and characterization of its special structure are essential to better understand the population's interaction with its environment and can be used to support future decisions.

It was evident that the ACP model presented a better fit of the data compared to the Poisson model. From the results of the ACP model, it was possible to predict when and where the bronchiolitis hospitalizations are distributed. The months in which the disease occurred most were identified according to each health district in Paraná State.

Although some maps were created to improve the visualization of spatial variability, some analysis, such as Moran Indices, can be performed in future analysis.

Furthermore, it is important to highlight that we analyzed count data instead of rates. The advantage is that we do not need to wait for data on live births to become available and the results and analyses can be more up to date. However, analyzing only the number of occurrences hampers the comparison among different regions as regions with more live births tend to have more disease cases.

**Acknowledgments:** The authors acknowledge the financial support of CAPES.

**Conflict of Interest:** The authors declare that there is no conflict of interest.

## REFERENCES

1. Nasri F. O envelhecimento populacional no Brasil. *Einstein*. 2008; 6(1): S4-6.
2. Piedimonte G, Perez MK. Respiratory Syncytial Virus Infection and Bronchiolitis. *Pediatr Rev*. 2014;35(12):519-30.
3. Caballero MT, Polack FP. Respiratory syncytial virus is an "opportunistic" killer. *Pediatr Pulmonol*. 2018;53(5):664-7.
4. Thorburn K, Fulton C, King C, Ramaneswaran C, Alammar A; McNamara PS. Transaminase levels reflect disease severity in children ventilated for respiratory syncytial virus (RSV) bronchiolitis. *Sci Rep*. 2018;8(1803).
5. Andrade DO, Botelho C, da Silva JLR, Faria SS, Rabahi MF. Sazonalidade climática e hospitalizações em crianças menores de cinco anos com doença respiratória. *Hygeia*. 2015; 11(20):99.
6. Fokianos K, Kedem B. Partial likelihood inference for time series following generalized linear models. *J Time Ser Anal*. 2004;25(2):173-97.
7. Cameron AC, Trivedi PK. *Regression Analysis of Count Data*. 2th ed. Cambridge University Press; 2013. 489 p.
8. Heinen A. Modelling time series count data: An autoregressive conditional Poisson model. *MPRA*. 2003;8113:1-37.
9. Gneiting T, Raftery AE. Strictly proper scoring rules, prediction, and estimation. *J Am Stat Assoc*. 2007;102(477):359-78.
10. Czado C, Gneiting T, Held L. Predictive model assessment for count data. *Biometrics*. 2009;65(4):1254-61.
11. Bush A, Thomson H. Acute bronchiolitis. *BMJ*. 2007;335(7628):1037.
12. Caillère N, Josseran L. Surveillance and epidemiology of infant bronchiolitis in France. *Arch Pediatr*. 2008;15(3):327-8.