PUBLICAÇÃO OFICIAL DO NÚCLEO HOSPITALAR DE EPIDEMIOLOGIA DO HOSPITAL SANTA CRUZ E PROGRAMA DE PÓS GRADUAÇÃO EM PROMOÇÃO DA SAÚDE - DEPARTAMENTO DE BIOLOGIA E FARMÁCIA DA UNISC

Revista de Epidemiologia e Controle de Infecção

Hospital Santa Cruz

ORIGINAL ARTICLE

Statistical modeling of hospital admissions for pneumonia in Campo Grande

Modelagem estatística das internações hospitalares por pneumonia em Campo Grande

Modelado estadístico de ingresos hospitalarios por neumonía en Campo Grande

https://doi.org/10.17058/reci.v11i3.15517

Received: *02/09/2020*
Accepted: *14/01/2021*
Available online: *20/01/2022*

Corresponding Author:
Amaury de Souza
amaury.de@uol.com.br

Universidade Fderal de Mato Grosso do Sul, caixa postal 549, CEP-79090-900. Campo Grande, MS.

Amaury de Souza[1];
Debora Aparecida Silva Santos[2];
Marcel Carvalho Abreu[3];
José Francisco de Oliveira-Júnior[4];
Pedro Luiz Ramos[5];
Francisco Louzada[6];
Elania Barros da Silva[7];
Pobocikova Ivana[8];

[1] *Universidade Federal de Mato Grosso do Sul, MS, Brasil.*
[2] *Universidade Federal de Mato Grosso, MT, Brasil.*
[3] *Universidade Federal Rural do Rio de Janeiro, Rio de Janeiro, Brasil.*
[4] *Universidade Federal de Alagoas, Maceió, Brasil.*
[5] *Pontifícia Universidade Católica do Chile Santiago, Chile.*
[6] *Universidade de São Paulo, São Carlos, Brasil*
[7] *Universidade Federal Fluminense, Niterói, Brasil.*
[8] *University of Žilina, Žilina, Slovakia.*

ABSTRACT

Justification and Objectives: Brazil lacks consistent epidemiological data on the respiratory morbidity of children and older adults, which makes it difficult to plan and execute effective preventive and health promotion actions. The objective of this study was to analyze the adjustments of distributions (Weibull, Normal, Gamma, Logistic) of historical series of hospitalizations for respiratory diseases (total hospitalizations), from 2011 to 2015, in Campo Grande, Mato Grosso do Sul. **Methods:** to determine the statistical models, four statistical indicators (coefficient of determination, mean root square error, mean absolute error and mean absolute percentage error) were performed from 2011 to 2015. Parameter estimates are obtained for the models adopted in the study, with and without a regression structure. **Results:** the results showed that Weibull, Gamma, Normal and Logistic distributions, applied to the series of hospitalizations for respiratory diseases in Campo Grande, were satisfactory in determining the shape and scale parameters, and the statistical indicators R^2, MAE, RSME and MAPE confirmed the data goodness-of-fit, and the graphical analysis indicated a satisfactory distribution fit. **Conclusion**: the analysis of monthly values indicates that Gamma is the best of the four distributions based on those selected. The regression model can be adjusted to the data and used as an alternative distribution that describes the hospitalization data considered in Campo Grande, Brazil.

Keywords: *Environmental Statistics. Infections. Periodicity. Probability. Hospitalization.*

RESUMO

Justificativa e Objetivos: o Brasil carece de dados epidemiológicos consistentes sobre a morbidade respiratória de crianças e idosos, o que dificulta o planejamento e a execução de ações efetivas de prevenção e promoção da saúde. O objetivo deste estudo foi analisar os ajustes das distribuições (Weibull, Normal, Gamma, Logística) da

Please cite this article as: *Souza, A. de, Abreu, M. C., de Oliveira-Júnior, J. F., Ramos, P. L., Santos, D. A. S., Louzada, F., da Silva, E. B., & Ivana, P. . (2022). Modelagem estatística das internações hospitalares por pneumonia em Campo Grande . Revista De Epidemiologia E Controle De Infecção, 11(3). https://doi.org/10.17058/reci.v11i3.15517*

série histórica de internações por doenças respiratórias (total de internações), no período de 2011 a 2015, em Campo Grande, Mato Grosso do Sul. **Métodos**: para determinar os modelos estatísticos, foram executados quatro indicadores estatísticos (coeficiente de determinação, erro quadrático médio, erro absoluto médio e erro percentual absoluto médio) de 2011 a 2015. As estimativas dos parâmetros são obtidas para os modelos adotados no estudo com e sem uma estrutura de regressão. **Resultados:** os resultados mostraram que as distribuições Weibull, Gamma, Normal e Logística, aplicadas à série de internações por doenças respiratórias em Campo Grande, foram satisfatórias na determinação dos parâmetros de forma e escala, e os indicadores estatísticos R^2, MAE, RSME e MAPE confirmaram a qualidade do ajuste dos dados, e a análise gráfica apontou um ajuste satisfatório das distribuições. **Conclusão**: a análise dos valores mensais indica que a Gamma é a melhor das quatro distribuições baseadas nos selecionados. O modelo de regressão pode ser ajustado aos dados e ser usado como uma distribuição alternativa que descreve os dados de internação considerados em Campo Grande, Brasil.

Descritores*: Estatísticas Ambientais. Infecções. Periodicidade. Probabilidade. Internação.*

RESUMEN

Justificación y Objetivos: el Brasil carece de datos epidemiológicos consistentes sobre la morbilidad respiratoria de niños y ancianos, lo que dificulta la planificación y ejecución de acciones efectivas de prevención y promoción de la salud. El objetivo de este estudio fue analizar los ajustes de las distribuciones (Weibull, Normal, Gamma, Logística) de la serie histórica de hospitalizaciones por enfermedades respiratorias (hospitalizaciones totales), de 2011 a 2015, en Campo Grande, Mato Grosso do Sul. **Métodos:** para la determinación de los modelos estadísticos, se realizaron cuatro indicadores estadísticos (coeficiente de determinación, raíz del error cuadrático medio, error medio absoluto y error porcentual absoluto medio) de 2011 a 2015. Se obtienen estimaciones de los parámetros para los modelos adoptados en el estudio, con y sin estructura de regresión. **Resultados**: los resultados mostraron que las distribuciones Weibull, Gamma, Normal y Logística, aplicadas a la serie de internaciones por enfermedades respiratorias en Campo Grande, fueron satisfactorias en la determinación de los parámetros de forma y escala, y los indicadores estadísticos R^2, MAE, RSME y MAPE confirmaron la calidad de ajuste de los datos, y el análisis gráfico indicaron un ajuste satisfactorio de las distribuciones. **Conclusión**: el análisis de los valores mensuales indica que la Gamma es la mejor de las cuatro distribuciones en base a las seleccionadas. El modelo de regresión se puede ajustar a los datos y utilizar como una distribución alternativa que describe los datos de hospitalización considerados en Campo Grande, Brasil.

Palabras clave*: Estadística Ambiental. Infecciones. Periodicidad. Probabilidad. Hospitalización.*

INTRODUCTION

Brazil has a lack of consistent epidemiological data on the respiratory morbidity of children and older adults, which makes it difficult to plan and execute effective preventive and health promotion actions. It is known that age is a risk factor proportional for respiratory diseases, i.e., the younger/older, the greater the risk for the development of these diseases.[1]

Risk factors for hospitalization for respiratory diseases include: exposure to environmental pollutants, especially smoking; household agglomeration; deficit in nutritional status; climatic seasonality; incomplete immunization schedules; low socioeconomic status; and exposure to biological agents, such as pollen. Such factors mainly affect individuals at the extremes of age, such as children under 5 years old and older adults over 65 years old.[1]

Groups that are susceptible to respiratory diseases in children are highly susceptible to exposure to air pollutants. They have greater minute ventilation due to accelerated basal metabolism and greater physical activity when compared to adults, in addition to staying longer in outdoor environments. Based on body weight, the volume of air that passes through a child's respiratory tract at rest is double that of adults in similar conditions. Irritation by pollutants that would produce a weak response in adults can potentially result in significant obstruction in childhood. Moreover, a not yet fully developed immune system increases the possibility of respiratory infections. Older adults are susceptible to the adverse effects of exposure to air pollutants, because they have a less efficient immune system (immunosenescence), a progressive decline in lung function that can lead to airway obstruction and exercise limitation. There is a reduction in chest wall compliance and pulmonary hyperinflation, causing additional energy expenditure to perform respiratory movements, in addition to a functional decrease in organic systems.[1]

Several studies involving adjustment of estimates using probabilistic/theoretical models in relation to historical data series have been developed, highlighting the benefits in planning activities that minimize risks, among which precipitation, air temperature, solar radiation, concentration of polluting gases historical seires stand out. However, there is published work (according to the authors' best knowledge) on historical series of hospitalizations for respiratory diseases based on the methodology used in this research, developed by Sousa et al (2019)[2], which analyzed the adjustments of Burr (Bu), Inverse Gaussian 3P (IG3P), Lognormal (LN), Pert (Pe), Rayleigh 2P (Ra 2P) and Weibull 3P (W3P) distributions of series of hospitalizations for respiratory diseases (total admissions). Shape and scale parameters of distributions were determined, and, to check the observation data adjustment quality, goodness-of-fit tests (GOF) were applied, such as Kolmogorov-Smirnov (KS), Shapiro-Wilk (SW), Anderson-Darling (AD) and (χ^2) chi-square tests.

The objective of this study was to analyze the distribution adjustment (Weibull, Normal, Gamma, Logistic) of historical series of hospitalizations for respiratory diseases (total hospitalizations), from 2011 to 2015, in Campo Grande, Mato Grosso do Sul. The analysis was performed between people (children and adults) who contracted pneumonia. The parameter estimates related to the data set are obtained and using different models with and without a regression structure. Finally, the Gamma regression model is select using discrimination criteria, and the prediction of the mean number of admission and its respective 95% confidence intervals are obtained.

METHODS

In this study, a descriptive analysis of variables was performed. We used Weibull (W2), Gamma (G), Normal (N) and Logistic (L) functions to model hospital admission data in Campo Grande. Performance indicators are calculated by comparing observed values to predicted values. Observed values are the observation data classified values, while predicted values are the values obtained from the adjusted distribution.

Health data

The city of Campo Grande, state of Mato Grosso do Sul, (20° 27′16 "S, 54° 47′16" W, 650 m), is located on the plateau called Maracajú-Campo Grande, 150 miles from the beginning of the world's largest flood plain, the Pantanal (a natural region encompassing the world's largest tropical wetland area, and the world's largest flooded grasslands) (139 111 km^2), and an estimated population of 724,000 inhabitants. According to Souza et al (2019)[2], the climate in the region of Campo Grande has moderate temperatures ranging from 17.8°C, minimum, to 29.8°C, maximum, with an average of 22.7°C, with hot summer and well distributed rainfall, and average relative humidity is 72.8%.

To correlate the meteorological data with the aggravation of respiratory diseases, hospital admission data were collected, together with the SUS (Unified Health System) Department of Informatics (DATASUS).

Data available came from the SUS Hospital Information System (SIH/SUS), managed by the Ministry of Health, through the Health Care Secretariat, together with State Health Secretariats and the Municipal Health Department, and processed by DATASUS, of the Ministry of Health's Executive Secretariat.

Hospital units, SUS patients (public or private parties), send hospitalization information, provided through AIH (Hospital Admission Authorization), to municipal (full management) or state (the others) managers. This information is processed in DATASUS, generating credits related to services provided and forming a valuable database, which contains a large part of hospital admissions in Brazil.

It should be noted that SIH/SUS collects variables related to hospitalizations, such as patient identification and qualification, procedures, examinations and medical actions performed, diagnosis, reasons for discharge, amounts due, etc. Through the Internet, DATASUS provides the main information for tabulation on the SIH/SUS databases.

Study area

This is an ecological study of time series. This design is characterized by studying groups of individuals, generally affected by geographic regions. In the case of this work, the site studied is the city of Campo Grande, state of Mato Grosso do Sul.

The study population was made up of people living in Campo Grande who were hospitalized due to respiratory diseases, from 2011 to 2015. We analyzed all hospitalizations with diagnosis of respiratory diseases from all hospitals at SUS. These data are AIH records of public and private hospitals and that assist part of the population that does not have health insurance nor are private or funded patients by companies (http:/www.datasus.gov.br). The information in the database comprises: the number of the hospital's Taxpayer's Registry (CGC), the city where it is located, patient age, sex, cause of hospitalization, procedure performed, patient zip code, hospitalization, date of discharge or death, days of UTI stay, among other information. The information in the database that were selected for this study were date of hospitalization, diagnosis, and patient age (children 0-4 years and adults 4-50 years).

Probability distributions

In this study, the effectiveness of four one-component probability models was assessed, namely W2, N, G and L. The one-component parametric probability distribution functions (pdfs) were used because our data exhibit unimodal shape. These four models have been selected among other one-component models due to their successful applications according to literature.

Estimation of model parameters

Several methods can be used to estimate the considered model parameters[3]. However, the selection of effective models is more important compared to the selection of parameter estimation methods[4]. In this work, the Maximum Likelihood Estimation (MLE) method is applied. This method has shown good results in several studies. It gives the values of the parameters which maximize the probability of obtaining the observed data.

The likelihood function $L(\theta)$ for a random sample x_1, $x_{2,\dots,} x_n$ and theoretical probability density function f with k parameters $\theta = (\theta_1, \theta_2, \dots, \theta_k)$ is represented by:

(1)

$$L(\theta)=\prod_{i=1}^{n} f(x_i,\theta)$$

For each parameter - θ_i, MLE involves maximizing the likelihood function by solving the following:

(02)

$$\frac{d\log(L(\theta))}{d\theta_i}=0,\ i=1,2,\ldots,k$$

Accuracy tests

The performance and accuracy of the tested models were assessed based on various statistical indices. The statistical indicators used in this study can be regrouped into two groups: i) dispersion indicators (error indicators); and ii) general performance indicators [4]. Mean Bias Error (MBE), Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) are noted in the dispersion indices, while Coefficient of Determination (R^2) are regrouped in the performance indicators[4]. The equations (Eq. 3 to 7) for statistical indicators are listed below:

(03)

$$MBE=\frac{1}{n}\sum_{i=1}^{n}(P_i-O_i)$$

(04)

$$RMSE=\sqrt{\sum_{i=1}^{n}\frac{(P_i-O_i)^2}{n}}$$

(05)

$$MAPE=\sum_{i=1}^{n}\frac{|P_i-Oi|}{P_i}\cdot\frac{100}{n}$$

(6) where P_i is the predicted value, O_i is the observed value, is the mean of the predicted values, is the mean of the observed values, and n is the number of observations.

$$R^2=\left[\frac{\sum_{i=1}^{n}(O_i-\bar{O})\cdot(P_i-\bar{P})}{\sqrt{\sum_{i=1}^{n}(O_i-\bar{O})^2\cdot\sum_{i=1}^{n}(P_i-\bar{P})}}\right]$$

Akaike Information Criterion

An information criterion that has been widely used in model selection is the Akaike Information Criterion (*AIC*). The *AIC* is defined according to the following Eq. (8):

(8)

$$AIC=-2\ln(L(\theta))+2k$$

Where *L*(θ) is the likelihood function at the maximum point of the model, and *k* is the number of parameters considered from said model. The first term of the equation is a reward for the quality of fit, and the second is a penalty for increasing the number of model parameters. The preferred model will be the one with the lowest *AIC*.

RESULTS

Table 1 shows the descriptive analysis of hospitalizations (morbidity) for respiratory diseases (pneumonia) for children and adults with daily averages for the months of 2011-2015. During the study period (January 1, 2011 to December 31, 2015), the number of hospitalizations for respiratory diseases was 609 (314 children and 295 adults, with a mean of 5 daily admissions, with a minimum of 2 and a maximum of 9).

Based on the data, it is important to check if there is a pattern in the proposed data set; therefore, we present in figure 1 the data set time series.

Table 1. Hospital admissions (morbidity) due to respiratory diseases (pneumonia) in children and adults in Campo Grande, from 2011 to 2015.

Children	**Jan**	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**
2011	5	6	4	9	8	7	9	7	8	8	8	6
2012	3	4	5	6	4	5	4	5	5	4	4	6
2013	4	3	4	4	4	5	5	7	4	5	6	4
2014	5	5	5	4	4	6	5	5	5	5	5	4
2015	2	3	3	4	5	6	6	6	7	6	6	4
Average	4	4	4	5	5	6	6	6	6	6	6	5
Stdev	1	1	1	2	2	1	2	1	2	2	1	1
Adults	**Jan**	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**
2011	4	5	3	7	6	5	7	5	6	6	6	5
2012	3	5	5	7	4	5	4	5	5	5	5	7
2013	4	3	4	4	4	4	4	7	4	5	6	4
2014	5	5	5	4	5	6	6	5	5	6	5	4
2015	2	3	3	4	5	6	5	6	7	6	6	4
Average	4	4	4	5	5	5	5	6	6	6	5	5
Stdev	1	1	1	1	1	1	1	1	1	1	1	1

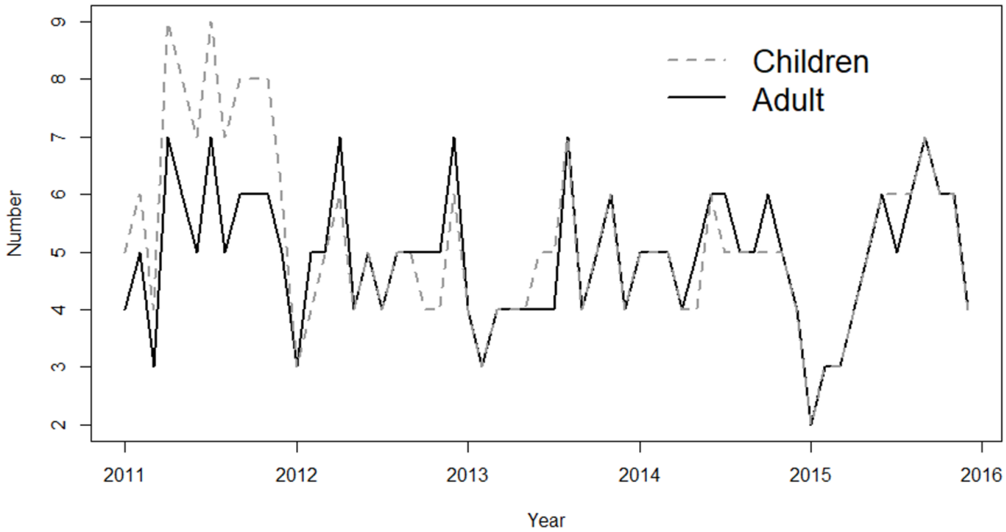


Figure 1. Time series for the monthly number of hospital admissions for 2011-2015.

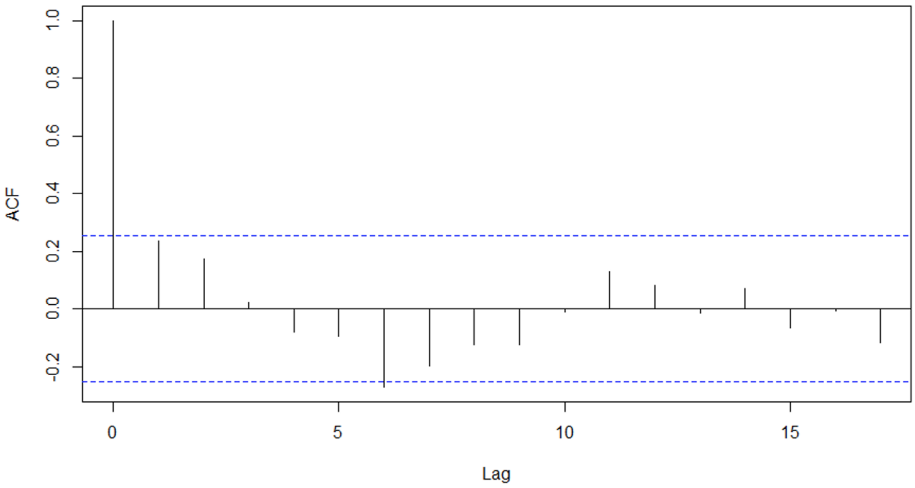


Figure 2. Autocorrelation for the monthly number of hospital admissions for the adult data.

Based on the data, a seasonal pattern was observed between rainy and dry seasons, and the transition period, especially in the quarters (April, May, June, July, August, and September), where the peak of hospitalizations corresponds to the dry season, low rainfall, relative humidity and minimum temperatures, and the period of highest burning rates in the state of Mato Grosso do Sul. Due to the small autocorrelation obtained from Figure 2, firstly, we will consider assumption that the data set is independent and identically distributed. Further, we will extended the analysis for using regression models where the covariates are related to dry or rainy seasons.

Model fitting with parameter estimation

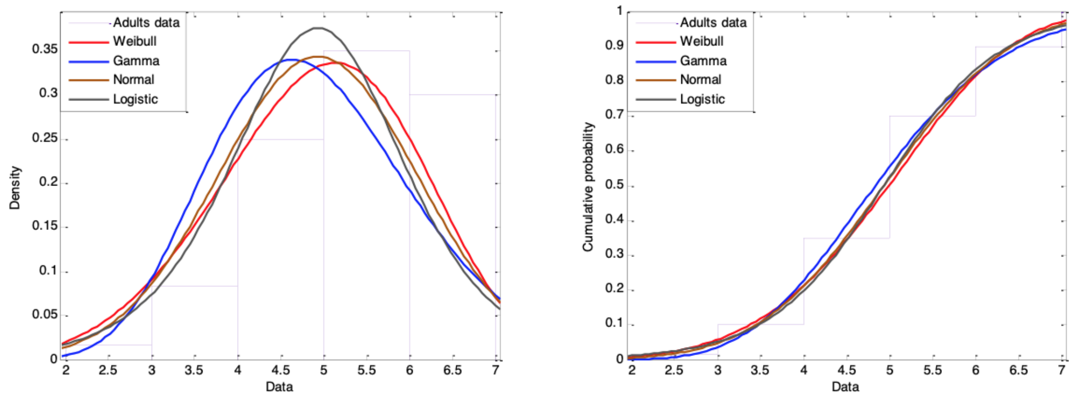
The tested distributions' parameters estimates were obtained using MLE and are presented in table 1. These parameters were computed using the statistical software R.

Please cite this article as: *Souza, A. de, Abreu, M. C., de Oliveira-Júnior, J. F., Ramos, P. L., Santos, D. A. S., Louzada, F., da Silva, E. B., & Ivana, P. . (2022). Modelagem estatística das internações hospitalares por pneumonia em Campo Grande . Revista De Epidemiologia E Controle De Infecção, 11(3). https://doi.org/10.17058/reci.v11i3.15517*

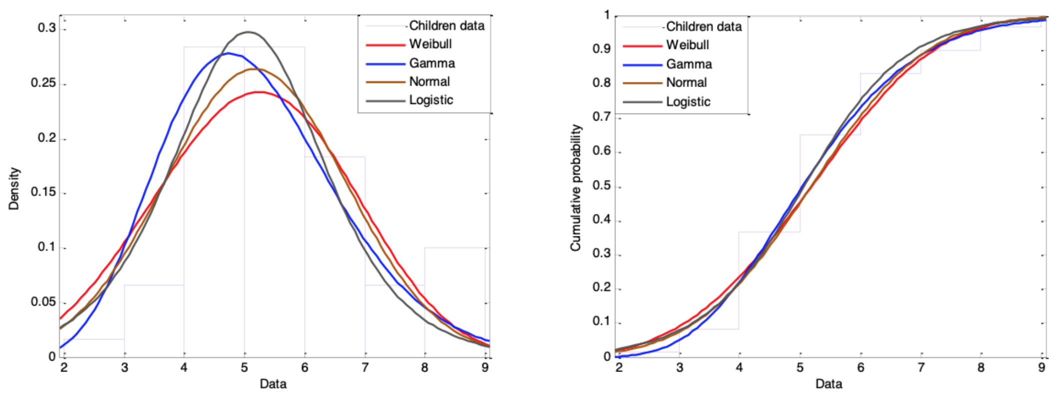
Table 2. Parameter estimates for the distributions studied related to the children and adult data.

Children	**θ**	**Estimates**		**R^2**	**MAE**	**RMSE**	**MAPE**
Weibull	a	5.740	pdf	0.5268	0.0509	0.0705	50.498
	b	3.631	cdf	0.9584	0.0646	0.0978	**17.751**
Logistic	μ	5.067	pdf	0.5803	0.0480	**0.0374**	51.330
	σ	0.839	cdf	0.9667	**0.0542**	**0.0575**	19.376
Gamma	a	12.132	pdf	**0.7148**	**0.0397**	0.0551	**48.721**
	b	0.427	cdf	**0.9754**	0.0618	0.0861	29.097
Normal	μ	5.183	pdf	0.5319	0.0480	0.0684	51.330
	σ	1.513	cdf	0.9589	0.0542	0.0848	19.376
Adult							
Weibull	a	5.385	pdf	0.6500	0.0607	0.0699	51.351
	b	4.813	cdf	0.9702	0.0832	0.1061	34.173
Logistic	μ	4.933	pdf	0.7198	0.0511	**0.0399**	35.696
	σ	0.665	cdf	0.9742	0.0817	**0.0625**	**25.676**
Gamma	a	16.781	pdf	**0.8677**	**0.0395**	0.0435	**29.351**
	b	0.294	cdf	**0.9877**	**0.0807**	0.0915	26.845
Normal	μ	4.933	pdf	0.7389	0.0528	0.0607	37.098
	σ	1.163	cdf	0.9769	0.0827	0.1010	25.892

The comparison of the four pdfs and their corresponding cdfs with the histogram of the monthly number of hospital admissions for 2011 to 2015 are illustrated in Figure 3.

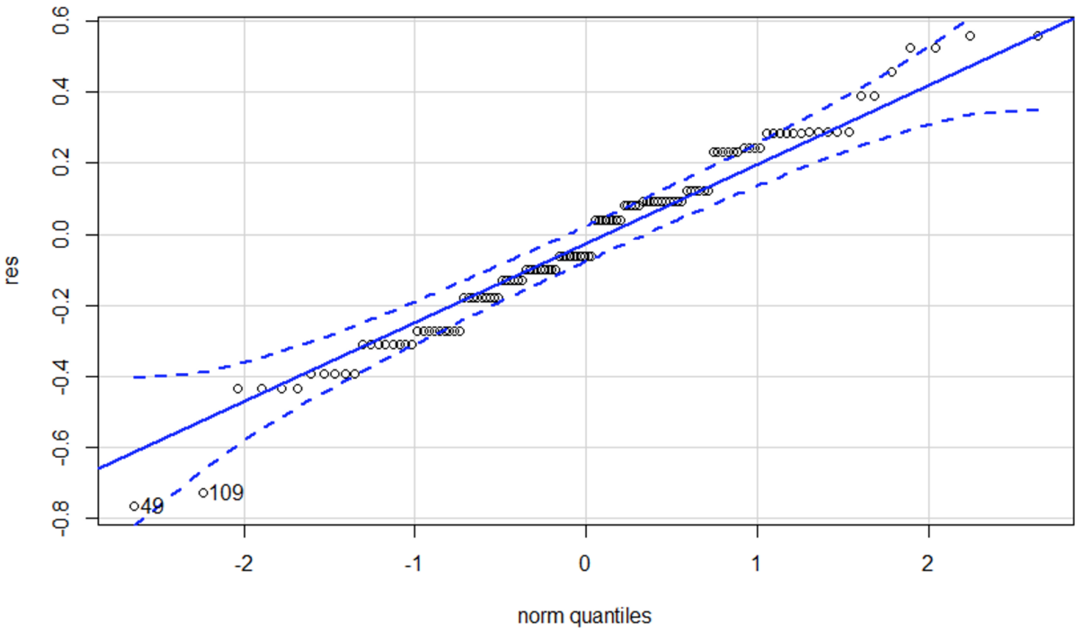


a) Adults



b) Children

Please cite this article as: *Souza, A. de, Abreu, M. C., de Oliveira-Júnior, J. F., Ramos, P. L., Santos, D. A. S., Louzada, F., da Silva, E. B., & Ivana, P. . (2022). Modelagem estatística das internações hospitalares por pneumonia em Campo Grande . Revista De Epidemiologia E Controle De Infecção, 11(3). https://doi.org/10.17058/reci.v11i3.15517*



c) Normal deviation chart with distribution for respiratory hospitalizations

Figure 3. a) pdf (left) and b) cdf (right) plots of the distributions obtained for the monthly number of hospital admissions for 2011-2015, c) normal deviation chart with distribution for respiratory hospitalizations in Campo Grande, for all age groups by period.

Table 2 show the results of the tests used. When choosing the best fitting model among the considered ones according to given criteria, we selected the model with lowest values of MAPE, MAE and RMSE, with R^2 closest to 1. The best values are highlighted in bold. According to the values of the model selection criteria, from the probability distributions considered, Gamma distribution satisfactorily fits to the hospital admission data for both age groups, followed by Logistic and Normal distributions. Gamma distribution is better than other distributions with respect to R^2; on the other hand, Logistic distribution achieves better values of RMSE. This is also evident and confirmed in the plots provided in figure 3.

So far, we have considered that the data set is i.i.d; however, we believe that dry and rainy season may increase or decrease the number of admissions. Therefore, we considered the same models in the presence of regression structure; the AIC values allow us to discriminate the models and are given in table 3.

By means of the selection criteria, we verified that, among the models used, Gamma distribution also returned the best fit for the proposed data. It should be noted that the selection criteria do not help to verify how well the model was adjusted, for this, residue analysis was considered to verify the model suitability. To achieve this, we apply KS and SW tests to verify the assumption of normal values of Gamma distribution's adjusted residuals.

In this case, if p-value is greater than 0.05, residuals follow a standard normal distribution and, therefore, the analysis is being performed adequately. Such a factor can also be confirmed by visual techniques such as the quantil-quantil plot given below:

We observe that almost all points are within the confidence interval of the adjusted curve, showing a good fit of Gamma distribution with covariates to describe the model.

In this way, with the regression model, we can easily make predictions for patient admission based on the information of the period of the year and whether they are adults or children (Table 3).

Table 3. AIC for used distributions, tests of normality of Gamma distribution's adjusted residuals, adjusted parameters for the Gamma regression model and prediction of the mean number of admission and its respective 95% confidence intervals.

							Type of patient	**Period**	**Estimates**	**95%CI**
Distributions	AIC	p-value	Coefficient	Estimated	Std. Error	p-value	Adult	Dry	5.3	(4.9; 5.7)
Weibull	411.11	0.43	Intercept	5.316	0.213	<0.0001		Rainy	4.6	(4.2; 4.9)
Gamma	404.87	0.10	Child / Adult	0.228	0.236	0.3360	Child	Dry	5.5	(5.1; 6.0)
Logistic	411.10		Period	-0.743	0.238	0.0023		Rainy	4.8	
Normal	406.86									(4.4; 5.2)

DISCUSSION

The incidence of respiratory diseases/pneumonia in children/adults is influenced by several factors. It appears that air pollution is the main risk factor related to these diseases, followed by natural climatic conditions and, to a lesser extent, viral infections, behavioral and/or domestic factors and family history of the disease. Likewise, it appears that the most investigated pollutant, among the filtered studies, is particulate material with a diameter of less than 2.5 micrometers (μm) - ($PM_{2.5}$), a type of inhalable particles, and it is shown that it is harmful even within the ideal limits established by World Health Organization (WHO).

Studies point to the relevance of exposure to air pollution in the development of respiratory diseases.[5-7] There is important evidence that exposure to pollutants, especially $PM_{2.5}$, corresponds to a strong risk factor for respiratory diseases. Above 50 μg/m³, for each addition of 10 μg/m³ in the levels of $PM_{2.5}$ there is a 2% increase in clinical visits, while, below this concentration, each increase in the same proportions resulted in an increase of 1% in the same variable.[8] In Brazilian cities, the increase of 3 to 5 μg/m³ was associated with an increase in the risk of developing respiratory diseases/pneumonia and in hospital admissions for children/adults, which varied according to the city.[5-7]

The impacts of air pollution on respiratory, cardiovascular and metabolic health have been studied and highlighted, since it is a risk factor caused by human activities, and can be reduced. However, it should be noted that, even at levels considered safe by WHO, air pollution may have contributed to the increase in the number of hospitalizations of children/adults for respiratory problems. Children are more susceptible to the effects of air pollution and, therefore, can develop acute respiratory symptoms more easily,[6] as they have greater lung ventilation than adults.[7] Still, children spend less time outdoors and, for this reason, have less contact with air pollution. However, the brief contact with pollution can have more severe effects.[7]

Inhalation of $PM_{2.5}$, due to the reduced aerodynamic size of its particles, can damage cardiopulmonary tissues, cross the alveolar-capillary barrier and reach the blood circulation, so that it can compromise other systems besides the respiratory system.[8] At the pulmonary level, it can directly lead to the formation of reactive oxygen species (ROS), activate macrophages and increase the production of inflammatory cytokines, propagating the inflammatory response.[9]

PM_{10} is less harmful than $PM_{2.5}$, due to the greater aerodynamic size of its particles (2.5-10 μm) that restrict its reach to the upper respiratory tract.[10] One of the main sources of this pollutant is burning, which is common in the central west region, in the dry season,[8,10] where the release of PM_{10} overlaps with that of $PM_{2.5}$.[10]

The levels of NO_2 and O_3 were also correlated with the risk of pneumonia incidence in children in Campo Grande.[11] The increase of 3 μg/m³, corresponding to the interquartile difference, in NO_x concentrations, generated percentage increases in the order of 14.8%, even though the concentrations did not exceed the limits considered acceptable.[5] NO_x represents all nitrogen oxides present in the atmosphere, but as NO is rapidly oxidized to NO_2, the NO_x concentration reflects the NO_2 concentration.[12] O_3 was also correlated with respiratory problems.[5] The increase in O_3 concentration by 10 μg/m³ was associated with an increase of 3.91% in respiratory diseases in children under 5 years of age.[5]

These pollutants are capable of increasing the bronchial epithelial permeability, consequently, leukocyte infiltration and the release of inflammatory mediators.[13] Considering that the pulmonary ventilation of children is greater than that of adults,[13] they become more prone to inflammation of the upper and lower airways due to exposure to air pollution and, therefore, represent a risk group for such effects. Chronic exposure to air pollution sensitizes the respiratory tract early, which is already inflamed by other health issues. Therefore, pollution represents a source of early and continuous allergic awareness.

Unlike atmospheric pollution by industrial and automobile sources, which are characterized by chronicity, the pollution resulting from fires is seasonal,[12,13] and is also related to the increase in the number of hospitalizations for respiratory diseases.[13,14] In addition to this, exposure to pollutants resulting from fires can exceed the places where they occur, due to air masses, leading to a higher incidence of respiratory diseases in other locations.[13] Exposure to air pollutants, regardless of the source, can often boost the effects of climatic factors on hospitalizations for respiratory events, and vice versa. In the Manaus region, hospital admissions appear to be more related to meteorological variables, especially humidity, than to pollutants from biomass burning.[11] There is evidence that there is a direct relationship between meteorological variations and the population's health, especially in the impairment of lung function and in the incidence of respiratory diseases.[11-14]

Variations in temperature and humidity are also related to changes in pollutant concentrations and, consequently, hospitalizations for respiratory diseases. The increase in temperature was associated with an increase in concentrations of $PM_{2.5}$.[15] Likewise, both the increase and the sudden reduction in temperature are related to the development of respiratory diseases, as well as the reduction of humidity and thermal comfort for the population.[10]

In the South region, most complaints of respiratory problems occurred in the coldest months (July-September),[16] and in the months that foresee the arrival of this season,[10] however, they were not correlated with precipitation rates.[16] In Rio de Janeiro, most complaints of viral respiratory infections occurred between late autumn and spring.[12]

Of the included studies, three verified the relationship between viral infections and children's respiratory problems. Viruses are largely responsible for the incidence of respiratory diseases in children. About 50%

of colds are of viral origin, and culminate in economic losses in relation to medical care, and social losses due to school absences.[16]

The higher incidence of respiratory diseases caused by viral causes in children is mainly due to their immature immune and respiratory system, which makes them more prone to infections and co-infections. Also, age interferes with the environment and individual behavior. At older ages, they spend most of the time in other places, due to curricular and extracurricular activities, which involve locomotion and contact with other risk factors. Specific environments and individual behavioral patterns can influence individual exposure to pollutants and clinical respiratory patterns.[13]

Behavioral and/or domestic factors (use of a wool blanket, smoking by parents, contact with household dust, among others) seem to have little influence on the children's respiratory, compared to the others mentioned. Although there is a mild effect, children with atopic allergy, renitis or asthma respond more severely to household dust.[16,17]

In this study, we list several risk factors for respiratory diseases in children/adults, mainly exposure to air pollutants, followed by natural climatic conditions, viral causes, behavioral and/or domestic factors, and family history of the disease. It is inferred that exposure to the association of different factors can aggravate and increase the incidence of respiratory diseases, and it is highlighted that factors such as pollution originate from human activities and, therefore, can be remedied. There is also the need to expand studies on the effects of exposure to pollutants in children/adults, due to the greater susceptibility of this age group.

Still, it was found that the pollutant most investigated in the studies and that is related to a greater chance of developing respiratory diseases in children is $PM_{2.5}$, and that it can cause damage even within the limits considered ideal. Therefore, the relationship between air pollution and hospital admissions can be extrapolated to small cities, where air quality is still in line with WHO parameters.

Understanding the main risk factors associated with respiratory diseases in children/adults is essential to plan public health actions, with a view to reducing and preventing the incidence of these diseases.

In this study, parameter values for all proposed models, Weibull, Normal, Gamma, Logistic, were estimated using MLE. Four different performance indicators were used to test the accuracy of the estimated distributions. Table 3 summarizes the results related to the MLE estimation method. Through a detailed assessment of these results, it can be observed that the Gamma model based on MLE obtained the best results using all indicator tests. Based on the RMSE results, the model recorded the lowest values among the other models. MAE values corroborate the RMSE results, which confirmed the model's superiority over the others. A closer inspection of accuracy analysis results shows that the results obtained with the R^2 test indicate a high degree of consistency between the observations and predicted pdfs, with a value of 0.9.

Souza et al.[2] adjusted the Burr (Bu), Inverse Gaussian 3P (IG3P), Lognormal (LN), Pert (Pe), Rayleigh 2P (Ra 2P) and Weibull 3P (W3P) distributions of the historical series of hospitalizations for diseases (total hospitalizations). According to the authors, the Gamma probability distribution function proved to be adequate to represent hospitalizations in all months of the year. Adjustments of probabilistic models can be seen in figures 3a and b. In general, a good fit of the models to the observed data is perceived. A proximity in the behavior of the curves of the different models with the parameters estimated by MLE obtained more indications as the most adequate pdf to represent the observed data (Table 3). These results corroborate those obtained by Souza et al.,[2] who verified compliance with the distributions to the monthly hospitalization data in Campo Grande.

To reduce uncertainties in hospital admission estimates, this study compared four candidate distributions (Weibull, Normal, Gamma, Logistic), in order to select the pdf that best matches hospital admissions dataset. The monthly data of hospital admissions, from 2011 to 2015, in Campo Grande, Brazil, were adjusted by the distributions considered. To determine the effectiveness of the statistical models, the four fit quality tests (R^2, RMSE, MAE and MAPE) were performed. The analysis of the monthly values indicates that Gamma is the best out of the four distributions based on the selected criteria.

Based on the results obtained, it can therefore be concluded that Gamma regression model fits the data and can be used as an alternative distribution that adequately describes hospital admission data considered in Campo Grande, Brazil.

FUNDING

This research received no external funding.

ACKNOWLEDGMENTS

The authors would like to thank their universities for their support.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest**.**

DATABASE DECLARATION/DATA AVAILABILITY

The climate database is in the public domain and is available at: https://www.cemtec.ms.gov.br/ and the hospital admissions database is available at http://www2.datasus.gov.br/DATASUS/index.php?area=02

ETHICAL CONSIDERATIONS

The present study is based on secondary, publicly available data, which do not constrain groups of populations and/or individuals in the presentation of the results found, ensuring the confidentiality of the information collected. Thus, the ethical aspects of research with human beings were respected, according to Resolution 466/2012.

REFERENCES

1. Shukla S.D. et al. Infection-Induced Oxidative Stress in Chronic Respiratory Diseases. In: Maurya P., Dua K. (eds) Role of Oxidative Stress in Pathophysiology of Diseases. Springer, Singapore. 2020. pp 125-147. doi: 10.1007/978-981-15-1568-2_8
2. Souza A, Aristone F, Fernandes WA, Olaofe Z, Abreu MC, Oliveira Júnior, JF, Cavazzana, G, Santos, CM. Statistical Behavior of Hospital Admissions for Respiratory Diseases by Probability Distribution Functions. J Infect Dis Epidemiol 2019, 5(6):098. doi: 10.23937/2474-3658/1510098
3. Tizgui I, El Guezar F, Bouzahir H, Benaid B. Comparison of Methods in Estimating Weibull Parameters for Wind Energy Applications. International Journal of Energy Sector Management 2017, 11(4): 650-663. doi: 10.1108/IJESM-06-2017-0002
4. Mohammadi K, Alavi O, McGowan JG. Use of Birnbaum-Saunders distribution for estimating wind speed and wind power probability distributions: A review. Energy Conversation and Management, 2017, 143: 109–122. doi: 10.1016/j.enconman.2017.03.083
5. Roman V V L, Carvalho Júnior, J A, Nascimento L F, Cesar A C G. Efeitos de poluentes do ar e doenças respiratórias utilizando dados estimados por modelo matemático. Revista Ambiente e Água 2015, 10(4): 825-831. doi: 10.4136/ambi-agua.1592
6. César A C G, Nascimento L F C, Mantovani K C C, Vieira L C P. Material particulado fino estimado por modelo matemático e internações por pneumonia e asma em crianças. Revista Paulista de Pediatria 2016, 34(1): 18-23. doi: 10.1016/j.rppede.2015.12.005
7. Nascimento A P S, Santos J, Mill J G, Souza J B, Reis Júnior N C, Reisen V A. Associação entre concentração de partículas finas na atmosfera e doenças respiratórias agudas em crianças. Revista de Saúde Pública 2017, 51(3): 1-10. doi: 10.1590/S1518-8787.2017051006523
8. Shen Y, Wu Y, Chen G, Van Grinsven H J M, Wang X, Gu B, Lou X. Non-linear increase of respiratory diseases and their costs under severe air pollution. Environmental Pollution 2017, 224(1): 631-637. doi: 10.1016/j.envpol.2017.02.047
9. Orona N S, Ferraro S A, Astort F, Morales C, Brittes F, Boero L, Tiscornia G, Maglione G A, Saldiva P H N, Yakisch S, Tasat D R. Acute exposure to Buenos Aires air particles (UAP-BA) induces local and systemic inflammatory response in middle-aged mice: A time course study. Environmental Pollution 2016, 208(A): 261-270. doi: 10.1016/j.envpol.2015.07.020
10. Ambiente M D M. Poluentes atmosféricos. 2018. Disponível em: http://www.mma.gov.br/cidadessustentaveis/qualidadedoar/poluentesatmosf%C3%A9ricos#Material_particulado.
11. Souza A, Ozonur D. Statistical Behavior of O3, OX, NO, NO2, and NOx in Urban Environment. Ozone-Science & Engineering The Journal of the International Ozone Association 2019, 42(1): 1-13. doi: 10.1080/01919512.2019.1602468
12. Souza A, Aristone F, Fernandes W A, Olaofe Z, Oliveira A P G, Carvalho M A, Oliveira-Junior J F, Cavazzana G, Santos C M, Soares D G. Analysis of Ozone Concentrations Using Probability Distributions. Ozone-Science & Engineering The Journal of the International Ozone Association 2020, 1-12. doi: 10.1080/01919512.2020.1736987
13. César A C G, Carvalho Jr J A, Nascimento L F C. Association between NOx exposure and deaths caused by respiratory diseases in a medium-sized Brazilian city. Brazilian Journal of Medical and Biological Research 2015, 48(12): 1130–1135. doi: 10.1590/1414-431x20154396
14. Carvalho E K M A, Dantas R T, Carvalho J R M. Análise da influência entre as variáveis meteorológicas e doenças respiratórias na cidade de Campina Grande, PB. Revista Brasileira de Climatologia 2016, 18(12): 63- 79. doi: 10.5380/abclima.v18i0.41099
15. César A C G, Nascimento L F C, Mantovani K C C, Vieira L C P. Material particulado fino estimado por modelo matemático e internações por pneumonia e asma em crianças. Revista Paulista de Pediatria 2016, 34(1): 18-23. doi: 10.1016/j.rppede.2015.12.005
16. Leotte J, Trombetta H, Faggion H Z, Almeida B M, Nogueira M B, Vidal L R, Rabon S M. Impact and seasonality of human rhinovirus infection in hospitalized patients for two consecutive years. Jornal de Pediatria 2017, 93(3): 294-300. doi: 10.1016/j.jped.2016.07.004
17. Souza A, Santos D A S, Oliveira-Junior J F, Garcia A P O, Silva L B. Modeling of hospital admissions for respiratory diseases as a function of probability distribution functions. Research, Society and Development 2020, 9(8): 1-17. doi: 10.33448/rsd-v9i8.6501

AUTHORS' PARTICIPATION

All authors participated in the article preparation, review and writing, and data collection and analysis.