Contents Lists Available at: https://journal2.unusa.ac.id/

Medical Technology and Public Health Journal

Journal Home Page: https://journal2.unusa.ac.id/index.php/MTPHJ

Research Article

Forecasting Bronchopneumonia Disease Using the SARIMA Method: A Case Study at Hospital X

Mursyidul Ibad[1*] | *Fasha Aurora Syabana*[1]

[1]*Department of Public Health, Faculty of Health, Universitas Nahdlatul Ulama Surabaya, Indonesia*

***Corresponding Author**:
Mursyidul Ibad, Department of Public Health, Faculty of Health, Universitas Nahdlatul Ulama Surabaya, Indonesia.
Email: mursyidul.ibad@unusa.ac.id

***DOI:** 10.33086/mtphj.v9i1.7263*

Article History:
Received, January 21st, 2025
Revised, February 1st, 2025
Accepted, March 21st, 2025
Available Online: March 30nd, 2025

Please cite this article as:
Ibad, Mursyidul & Fasha Aurora Syabana, "Forecasting With SARIMA Method Bronchopneumonia Disease (Case Study At Hospital X)" Register: Medical Technology and Public Health Journal, Vol. 9, No. 1, pp. 109-117, 2025

ABSTRACT

Pneumonia remains the leading cause of death among children globally. In Indonesia, the prevalence of pneumonia in 2023 was particularly concerning, ranking first among diseases commonly affecting children under five, with a staggering 31.4% of cases. East Java Province reported the highest incidence, accounting for 32.03% or 45,041 affected children. To tackle this pressing issue, advancements in statistical methods, specifically the time series approach, offer valuable tools for prediction. This study employs the Seasonal Auto-Regressive Integrated Moving Average (SARIMA) forecasting method. Conducted as a non-reactive study utilizing secondary data, this research follows a retrospective cohort design, collecting data over a specific timeframe. The SARIMA (0,1,0)(1,1,0, 12) model is suitable for predicting bronchopneumonia cases, as it contains significant parameters and satisfies all necessary assumptions. The predictions generated from this model will extend over the following 12 periods.

Keywords: pneumonia, children, ARIMA, SARIMA

INTRODUCTION

Hospitals are essential healthcare institutions that offer a wide range of health services, including inpatient and outpatient care and emergency services (Sari, Adi, and Lestari, 2023). They play a crucial role in the social health system, delivering comprehensive care that includes curative disease treatment and preventive measures for the community (Santoso, 2024). The presence of adequate and high-quality inpatient facilities is a key indicator of successful hospital healthcare services.

Bronchopneumonia is characterized by inflammation of the lungs, affecting one or more lobes and marked by infiltrate patches caused by bacteria, viruses, fungi, or foreign objects. It is the most common clinical manifestation of pneumonia in children. Both bronchopneumonia and pneumonia involve lung inflammation. In pediatric patients, pneumonia impairs the lungs' ability to expand, prompting the body to breathe rapidly to prevent hypoxia. Symptoms of bronchopneumonia

may include chills, fever, pleuritic chest pain, a productive cough, nasal flaring, and accessory muscles during breathing (Dwiarindi and Naufal, 2022). The severity of this condition can vary from mild to severe, potentially leading to life-threatening complications (Ministry of Health, 2022).

According to the World Health Organization (2022), pneumonia stands as the leading cause of death among children worldwide. In Indonesia, pneumonia was identified as the most prevalent illness affecting children under five in 2023, with a prevalence rate of 31.4%. East Java Province reported the highest incidence of pediatric pneumonia, accounting for 32.03% of cases, totaling 45,041 individuals (Ministry of Health, 2023). Furthermore, data from the Central Statistics Agency (BPS) of East Java indicated that Surabaya City had the highest number of pneumonia cases among the 38 districts and cities, recording 11,692 cases.

In 2023, pneumonia emerged as the most prevalent illness affecting children under the age of five in Indonesia, with a prevalence rate of 31.4%. East Java Province reported the highest number of pneumonia cases among children, accounting for 32.03% of the total, or 45,041 cases (Ministry of Health, 2023). The Central Bureau of Statistics (BPS) of East Java indicated that Surabaya City recorded the most significant pneumonia cases among the 38 regencies and cities, totaling 11,692 cases.

Hospital X is a private hospital in Surabaya that is crucial in providing healthcare services to the community. According to medical record data from 2021 to 2023, bronchopneumonia emerged as the most prevalent illness among inpatient cases. Following bronchopneumonia, gastroenteritis, and colitis of unspecified origin ranked second, with a total of 3,228 cases. Non-insulin-dependent diabetes mellitus with multiple complications was next, totaling 3,203 cases. Urinary tract infections (UTIs) and typhoid fever ranked fourth and fifth, with 3,130 and 3,096 cases, respectively.

To address this challenge, the time series approach is an advancement in statistical methods useful for prediction. The forecasting method selected for this study is the Seasonal Auto-Regressive Integrated Moving Average (SARIMA). This method was chosen due to its ability to capture time-based data patterns, particularly trends and seasonality, making it well-suited for predicting the number of bronchopneumonia cases. The SARIMA method consists of four main components: seasonal, Autoregressive (AR), Moving Average (MA), and Differencing (d), which are utilized to achieve data stationarity.

In a study conducted by Wibowo and Purwaningsih (2023), the SARIMA(2,1,0)(0,1,0) model produced a Mean Absolute Percentage Error (MAPE) value of 13.37% with an accuracy rate of 86.63%. This model was deemed adequate for forecasting visitor data at tourist attractions in Bantul Regency. Based on these insights, the researcher is interested in studying the forecasting of bronchopneumonia cases using the SARIMA method at Hospital X in Surabaya City.

MATERIAL AND METHODS

The research conducted is categorized as non-reactive and utilizes secondary data. The study's design is a retrospective cohort study, wherein data is collected over a specified period. The primary variables under investigation include the number of bronchopneumonia cases and the corresponding time, which are measured every month. This study leverages secondary data from the published records of inpatient cases at Hospital X. Specifically, the data comprises monthly time series records of bronchopneumonia cases from 2021 to 2024. The analysis employs a cohort data forecasting technique, utilizing the Seasonal Auto-Regressive Integrated Moving Average (SARIMA) method as the forecasting approach.

RESULTS AND DISCUSSION

Overview of Bronchopneumonia Case Numbers at Hospital X

This study examines data from a private hospital catering to general and insured patients. The reported numbers of bronchopneumonia cases are illustrated in Figure 1.

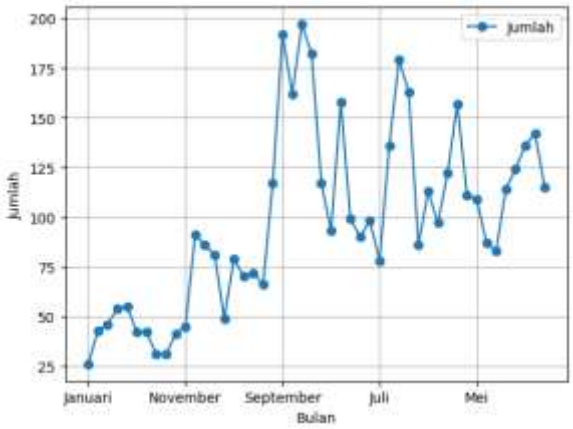

Figure 1. Plot of Bronchopneumonia Case Numbers at Hospital X from 2021 to 2024

As depicted in Figure 1, the count of bronchopneumonia cases at Hospital X from 2021 to 2024 exhibits variability, characterized by noticeable peaks and troughs. The highest recorded incidence of bronchopneumonia occurred in November 2022. Subsequently, the data reveals a consistent pattern of fluctuations, with increases and decreases occurring cyclically throughout the year.

Model Analysis and Forecasting Results of Bronchopneumonia Cases Using the ARIMA Method

Forecasting bronchopneumonia cases with the ARIMA method involves five key stages: stationarity testing, model identification, model estimation, diagnostic checks, and forecasting.

1. Stationary Testing

The initial step in data analysis is to assess the stationarity of the dataset concerning its mean and variance. The stationarity of the variance can be evaluated using the Augmented Dickey-Fuller (ADF) Test. Upon conducting the ADF test on bronchopneumonia case data from 2021 to 2024, the findings indicated that the data was not stationary at level 1. This step is illustrated in the following table:

Table 1. Results of the First Augmented Dickey-Fuller (ADF) Test

ADF Statistic	-2.791
p-value	0.06
Critical Values	1%: -3.578 5%: -2.925 10%: -2.601

According to Table 1, the initial Augmented Dickey-Fuller (ADF) test results indicated an ADF statistic value of -2.791. This value is less than the critical thresholds at the 1% significance level (-3.578) and the 5% level (-2.925). However, the resulting p-value is 0.06, which exceeds 0.05.

Consequently, the bronchopneumonia case data from Hospital X fails to satisfy the stationarity assumption, necessitating the data transformation.

Table 2. Results of the Second Augmented Dickey-Fuller (ADF) Test

ADF Statistic	-7.404
p-value	0.000
Critical Values	1%: -3.581 5%: -2.927 10%: -2.601

According to the findings presented in Table 2, the results from the second Augmented Dickey-Fuller (ADF) test revealed an ADF statistic of -7.404. This value is lower than the critical thresholds at the 1% significance level (-3.581), 5% (-2.927), and 10% (-2.601). Additionally, the p-value obtained is 0.00, which is less than 0.05. This finding indicates that the bronchopneumonia case data from Hospital X has satisfied the stationarity assumption following the first differencing process. We can refer to the ACF and PACF plots to assess stationarity regarding the mean. Below are the ACF and PACF plots for the bronchopneumonia case data from January to December 2024.

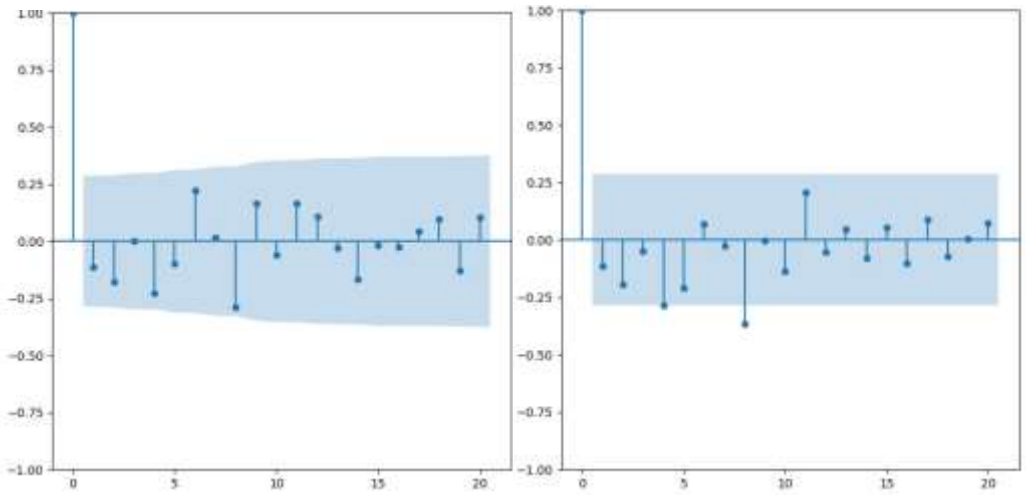

Figure 2. ACF and PACF Plot of Bronchopneumonia Case Data

Based on the ACF plot in Figure 2 and the PACF plot in Figure 3, it is observed that lags 1–17 fall within the blue boundaries. These blue boundaries signify the limits of significant autocorrelation. The data related to bronchopneumonia case numbers can be considered stationary concerning the mean if no lines cross the significant boundary by more than three lags. As such, we conclude that the data is stationary in terms of mean since no lag exceeds the blue boundary by more than this threshold.

2. Model Identification

The next step involves identifying a tentative SARIMA model based on previously obtained ACF and PACF plots. The bronchopneumonia case data has been rendered stationary in variance and mean through a single differencing process, resulting in a value of d = 1. Utilizing the insights from the ACF and PACF plots, the preliminary forecasting models for the bronchopneumonia case data are as follows:

SARIMA(0,1,1)(1,1,1)12, SARIMA(0,1,1)(1,1,2)12, SARIMA(0,1,1)(1,1,0)12, SARIMA(1,1,0)(1,1,0)12, SARIMA(0,1,1)(0,1,1)12, SARIMA(0,1,1)(2,1,1)12, SARIMA(0,1,0)(1,1,0)12, SARIMA(1,1,1)(1,1,0)12, SARIMA(1,1,1)(1,1,1)12, SARIMA(1,1,0)(0,1,1)12, SARIMA(2,1,2)(1,1,1)12, and SARIMA(0,1,1)(2,1,2)12.

3. Model Parameter Estimation

During the model parameter estimation stage, a function will be developed to perform SARIMA modeling, represented by the notation (p,d,q)(P, D, Q)s. These parameters are essential for capturing the seasonal patterns, trends, and noise present in the data. In this context, p refers to the non-seasonal autoregressive (AR) parameter, d denotes the non-seasonal differencing, and q represents the non-seasonal moving average (MA) parameter. Additionally, P signifies the seasonal AR component, D indicates the seasonal differencing, Q is the seasonal MA parameter, and s represents the period.

Table 3. SARIMA Model Parameter Estimation

No	Sarima Model	Parameter	*Coefisien*	P-value	Noted
1	$(0,1,1)(1,1,1)^{12}$	MA 1	-0.27	0.10	Not significant
		SAR 12	-0.06	0.92	Not significant
		SMA 12	-0.99	0.99	Not significant
2	$(0,1,1)(1,1,2)^{12}$	MA 1	-0.27	0.102	Not significant
		SAR 12	-0.62	0.99	Not significant
		SMA 24	0.97	0.97	Not significant
		SMA 12	-1.90	0.99	Not significant
3	$(0,1,1)(1,1,0)^{12}$	MA 1	-0.24	0.17	Not significant
		SAR 12	-0.55	0.01	Significant
4	$(1,1,0)(1,1,0)^{12}$	AR 1	-0.15	0.36	Not significant
		SAR 12	-0.55	0.00	Not significant
5	$(0,1,1)(0,1,1)^{12}$	MA 1	-0.27	0.10	Not significant
		SMA 12	-0.99	0.989	Not significant
6	$(0,1,1)(2,1,1)^{12}$	MA 1	-0.30	0.08	Not significant
		SAR 24	-0.82	0.00	Significant
		SMA 12	0.96	0.03	Significant
		SAR 12	-1.09	0.00	Significant
7	$(0,1,0)(1,1,0)^{12}$	SAR 12	-0.55	0.01	Significant
8	$(1,1,1)(1,1,0)^{12}$	AR 1	0.53	0.12	Not significant
		MA 1	-0.81	0.01	Significant
		SAR 12	-0.53	0.02	Significant

No	Sarima Model	Parameter	*Coefisien*	P-value	Noted
9	$(1,1,1)(1,1,1)^{12}$	AR 1	0.50	0.11	Significant
		MA 1	-0.80	0.00	Significant
		SAR 12	-0.05	0.94	Not significant
		SMA 12	-0.99	0.98	Not significant
10	$(1,1,0)(0,1,1)^{12}$	AR 1	-0.17	0.30	Not significant
		SMA 12	-0.99	0.99	Not significant
11	$(0,1,1)(2,1,2)^{12}$	MA 1	-0.30	0.08	Not significant
		SAR 12	-0.91	0.99	Not significant
		SAR 24	-0.48	0.99	Not significant
		SMA 12	0.11	1.00	Not significant
		SMA 24	-0.85	0.99	Not significant
12	$(2,1,2)(1,1,1)^{12}$	AR 2	-0.51	0.23	Not significant
		AR 1	-0.39	0.36	Not significant
		MA 1	0.21	0.99	Not significant
		MA 2	0.98	0.84	Not significant
		SAR 12	-0.41	0.32	Not significant
		SMA 12	-0.15	0.78	Not significant

In the model parameter estimation stage, the determination of significance can be made by observing the p-value for AR, MA, and the constant. The tentative models that pass the significance test are SARIMA(0,1,1)(1,1,0)12, SARIMA(0,1,0)(1,1,0)12, SARIMA(1,1,1)(1,1,0)12, and SARIMA(0,1,1)(2,1,1)12. These models have at least one parameter with a p-value less than α 0.05, indicating they are significant. Consequently, these four models are tentative and will be evaluated further in the model diagnostics stage.

4. Data Diagnostic

Once the tentative models have been selected, the subsequent step is to conduct model diagnostics. This diagnostic assessment aims to confirm whether the chosen models satisfy the underlying assumptions. The data diagnostics encompass a white noise test and a normality test. The p-value from the Ljung-Box test is analyzed to assess white noise within the model. The results of the Ljung-Box test are presented below:

Tabel 4. *Ljung* Box in Residual Test

No	Sarima Model	*Ljung-Box Lag 12*			Noted
		p-value	AIC	BIC	
1.	$(0,1,1)(1,1,0)^{12}$	0.130	362.555	367.221	*W*hite *Noise*

2.	$(0,1,0)(1,1,0)^{12}$	0.156	361.800	364.911	*White Noise*
3.	$(1,1,1)(1,1,0)^{12}$	0.255	362.365	368.586	*White Noise*
4.	$(0,1,1)(2,1,1)^{12}$	0.263	362.395	370.172	*White Noise*

Residuals can be considered white noise when the p-value exceeds 0.05. According to Table 5.4, the results of the Ljung-Box test applied to the residuals of the four SARIMA models indicate that all models have p-values greater than 0.05. This finding suggests that there is no significant autocorrelation present in the residuals. As a result, the residuals from all four models can be classified as white noise or random, indicating that the models have effectively captured the underlying patterns in the data.

Furthermore, the SARIMA(0,1,0)(1,1,0)12 model has the lowest AIC and BIC values among the models analyzed, with an AIC of 361.800 and a BIC of 364.911. This finding suggests that this model achieves the optimal balance between accuracy and complexity. Given the outcomes of the residual tests and the model selection criteria (AIC and BIC), the SARIMA(0,1,0)(1,1,0)12 model is identified as the best option. The next step will involve conducting a normality test using the Shapiro-Wilk test, with the results illustrated in the figure below.

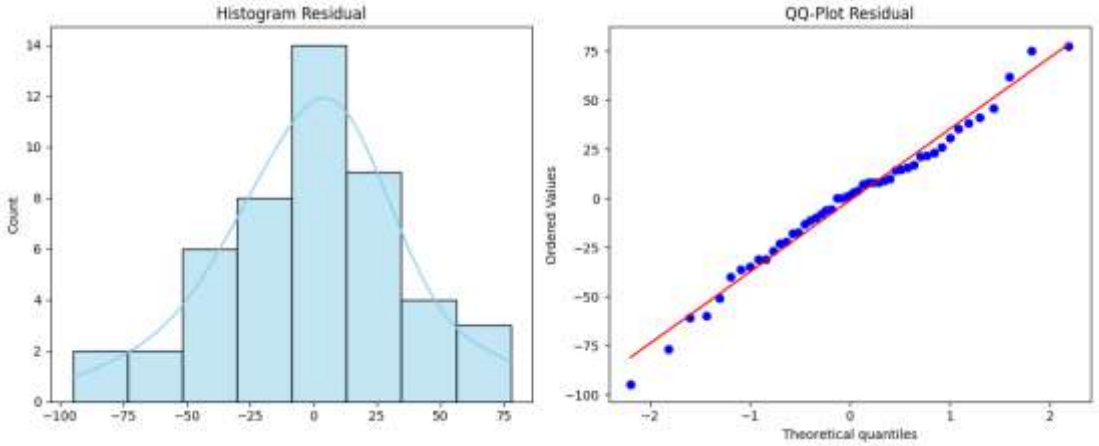

Figure 3. Residual Result

The results indicated that the SARIMA$(0,1,0)(1,1,0)_{12}$ model demonstrates a data distribution consistent with the standard line, suggesting that the data adheres to a normal distribution. Additionally, as shown in Table 5.5, the p-value is 0.818, which exceeds the 0.05 threshold, indicating that the SARIMA model satisfies the standard distribution assumption. Furthermore, the SARIMA$(0,1,0)(1,1,0)_{12}$ model can be mathematically expressed to derive an equation that will forecast the number of bronchopneumonia cases.

Table 5. Normality Test on Residuals

	Statistik	p-*value*
Shapiro-Wilk Test	0.9857	0.8180
Jarque-Bera Test	0.5627	0.7548

5. Forecasting Result

The forecasted number of bronchopneumonia cases using the SARIMA(0,1,0)(1,1,0,12) method is shown in Table 7.

Table 7. Forecasted Number of Bronchopneumonia Cases Using the SARIMA(0,1,0)(1,1,0,12) Method

Period	Forecast	Low	High
January 2025	71.99	52.40	91.60
February 2025	69.99	50.40	89.60
March 2025	121.52	101.93	141.13
April 2025	68.36	48.76	87.96
May 2025	62.50	42.91	82.11
June 2025	57.04	37.44	76.64
July 2025	44.22	24.62	63.82
August 2025	90.10	70.50	109.70
September 2025	118.28	98.68	137.88
October 2025	114.86	95.25	134.46
November 2025	75.12	55.52	94.72
December 2025	77.88	58.28	97.48

The SARIMA(0,1,0)(1,1,0,12) model forecasting indicates that the number of bronchopneumonia cases at Hospital X varies, showing monthly increases and decreases. This fluctuation is illustrated in the graph below:

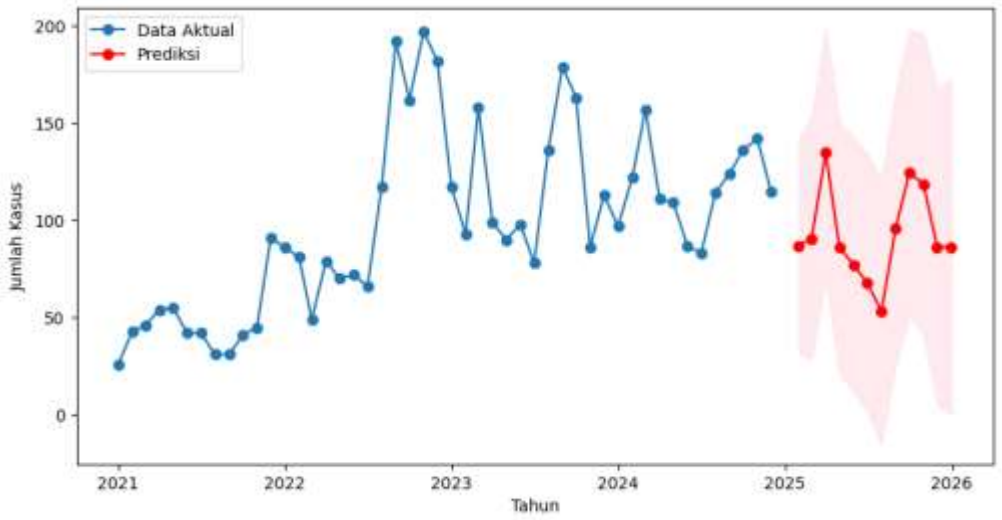

Figure 4.

Based on the analysis conducted using the SARIMA (0,1,0)(1,1,0)[12] model, the forecasted number of bronchopneumonia cases for the year 2025 is presented in Table 7. The forecasting results reveal monthly fluctuations, with the highest number of cases predicted in March (119.76 cases), September (117.92 cases), and October (113.77 cases). In contrast, the lowest numbers are anticipated in July (42.30 cases) and June (55.53 cases), indicating a discernible annual seasonal pattern.

CONCLUSION AND SUGGESTION

The analysis concludes that the SARIMA(0,1,0)(1,1,0,12) model is well-suited for predicting bronchopneumonia case data. The model's parameters are significant and fulfill all necessary assumptions. Predictions were generated for the upcoming 12 periods utilizing this model.

ACKNOWLEDGEMENT

Thank you to Universitas Nahdlatul Ulama Surabaya for granting permission to conduct this research, which facilitated the publication of this paper.

REFERENCES

Badan Pusat Statistik Jawa Timur. (2023). Jumlah Jenis Penyakit Malaria, TB Paru, Pneumonia, Kusta Menurut Kabupaten/Kota Di Provinsi Jawa Timur, 2023. Available at: https://jatim.bps.go.id/id/statistics-table/1/MzAwMSMx/jumlah-jenis-penyakit-malaria--tb-paru--pneumonia--kusta-menurut-kabupaten-kota-di-provinsi-jawa-timur--2022.html.

Dwiarindi, B.W. and Naufal, A. F. (2022) 'Penatalaksanaan Fisioterapi Pada Kasus Bronchopneumonia (a Case Report)', Journal of Innovation Research and Knowledge, 2(7), pp. 2825–2828. doi: https://bajangjournal.com/index.php/JIRK/article/view/4201.

Ministry of Health. (2023) Rencana Aksi Nasional Penanggulangan Pneumonia dan Diare 2023-2030. Available at: https://p2p.kemkes.go.id/wp-content/uploads/2023/12/NAPPD_2023-2030-compressed.pdf.

Santoso, H. (2024) 'Hak Rumah Sakit untuk Bekerjasama dengan Badan Penyelenggara Jaminan Sosial Kesehatan dalam Pelayanan Kesehatan', Jurnal Hukum Indonesia, 3(1), pp. 29–38. doi: 10.58344/jhi.v3i1.675.

Sari, S.I., Adi, R.P. and Lestari, S. (2023) 'Perancangan Enterprise Architecture Dengan Framework Togaf ADM Pada Rumah Sakit Umum (Studi Kasus: RSUD Abdul Moeloek Provinsi Lampung)', Z.A. Pagar Alam, pp. 339–350.

World Health Organization. (2022) 'Report: Pneumonia in children'. Geneva: WHO

Wibowo, R, Purwaningsih (2023) 'Asuhan Keperawatan Medikal Bedah pada An. F dengan penyakit Bronchopneumonia', Jurnal Keperawatan Medika, 2(1), pp. 38–42. doi: :https://jkem.ppj.unp.ac.id/index.php/jkem