Dai et al. BMC Public Health (2024) 24:2549
https://doi.org/10.1186/s12889-024-20033-7

BMC Public Health

RESEARCH

Open Access

Epidemiology and SARIMA model of deaths in a tertiary comprehensive hospital in Hangzhou from 2015 to 2022

Jingyuan Dai[1], Yun Xiao[1], Qionglian Sheng[1], Jing Zhou[2], Zhe Zhang[2] and Fenglong Zhu[3*]

Abstract

Background By analysing the deaths of inpatients in a tertiary hospital in Hangzhou, this study aimed to understand the epidemiological distribution characteristics and the composition of the causes of death. Additionally, this study aimed to predict the changing trend in the number of deaths, providing valuable insights for hospitals to formulate relevant strategies and measures aimed at reducing mortality rates.

Methods In this study, data on inpatient mortality at a tertiary hospital in Hangzhou from 2015 to 2022 were obtained via the population information registration system of the Chinese Center for Disease Control and Prevention. The death data of inpatients were described and analysed through a retrospective study. Excel 2016 was utilized for data sorting, and SPSS 22.0 software was employed for data analysis. The statistical inference of single factor differences was conducted via χ2 tests. The SARIMA model was established via the forecast, aTSA, and tseries software packages (version 4.3.0) to forecast future changes in the number of deaths.

Results A total of 1938 inpatients died at the tertiary hospital in Hangzhou, with the greatest number of deaths occurring in 2022 (262, 13.52%). The sex ratio was 2.22:1, and there were significant differences between sexes in terms of age, marital status, educational level, and place of residence ($P<0.05$). The percentage of males in the groups aged of 20 to 29 and 30 to 39 years was significantly greater than that of females ($\chi^2=46.905$, $P<0.001$). More females than males died in the widowed group, and divorced and married males experienced a greater number of deaths than divorced and married females did ($\chi^2=61.130$, $P<0.001$). The proportions of male students with a junior college and senior high school education were significantly greater than that of female students ($\chi^2=12.310$, $P<0.05$). The primary causes of mortality within the hospital setting included circulatory system diseases, injury, poisoning, tumours, and respiratory system diseases. These leading factors accounted for 86.12% of all recorded deaths. Finally, the SARIMA $(2, 1, 1)(1, 1, 1)_{12}$ model was determined to be the optimal model, with an AIC of 380.23, a BIC of 392.79, and an AICc of 381.81. The MAPE was 14.99%, indicating a satisfactory overall fit of this model. The relative error between the predicted and actual number of deaths in 2022 was 8.02%. Therefore, the SARIMA $(2, 1, 1)(1, 1, 1)_{12}$ model demonstrates good predictive performance.

*Correspondence:
Fenglong Zhu
906429443@qq.com

Full list of author information is available at the end of the article

Conclusions Hospitals should enhance the management of sudden cardiac death, acute myocardial infarction, severe craniocerebral injury, lung cancer, and lung infection to reduce the mortality rate. The SARIMA model can be employed for predicting the number of deaths.

Keywords SARIMA model, Epidemiology, Death cases, Hospital

Background

With the rapid development of the social economy and changes in the living environment, there has been a corresponding shift in the causes of death among residents [1]. The study of hospital fatalities can be specifically aimed at enhancing the standard of medical technology within hospitals and providing improved care for patients. Our hospital is a comprehensive three-level facility located in Hangzhou city, that has 1,012 open beds and 45 clinical medical technology disciplines. This makes our hospital a significant representative within the field. By using the population information registration system of the Chinese Center for Disease Control and Prevention, data on 1,938 hospitalized patients who died between 2015 and 2022 were meticulously collected. These data were analysed to investigate the epidemiological characteristics of these deaths.

In a comparative analysis of mortality patterns across various hospitals, there appeared to be some variation in the ranking of the leading causes of death. A statistical examination of mortality cases at a tertiary hospital in Sichuan Province revealed that malignant tumours, cardiovascular diseases, and respiratory diseases were the predominant causes of death [2], which was consistent with the results of Li Y et al. [3] However, a study conducted by Du J et al. [4] revealed that circulatory system diseases were the leading cause of death, followed by tumours, injury‒poisoning, and respiratory system diseases, among 3682 deaths in a hospital in Beijing. The results were different from those in Zhuhai, Guangdong [5] and Xi 'an, Shanxi [6]. Therefore, investigating the primary causes of mortality in the hospital's vicinity and enhancing the level of medical technology on the basis of specific circumstances are essential. This will ultimately lead to a reduction in disease-related fatality rates.

Currently, the autoregressive integrated moving average (ARIMA) model is extensively utilized in disease prediction, health expenditure prediction, weather forecasting, and other fields. Building upon the traditional ARIMA model, a new SARIMA model was constructed by incorporating seasonal factors [7–9]. The SARIMA model is capable of capturing the periodicity, trend, and randomness of data, and it has been shown to have higher prediction accuracy than the ARIMA model. James A [10] utilized the SARIMA model to predict the death toll of the novel coronavirus in Brazil. On the other hand, Qi Feng [11] et al. utilized the number of cancer deaths attributed to smoking in Qingdao to develop a SARIMA model for predicting the trend of smoking-related cancer fatalities. Furthermore, Guo J [12] et al. applied the SARIMA model to analyse meteorological data from the Hailun Agricultural Ecology Experimental Station. However, there is a scarcity of published research on the application of SARIMA models in predicting hospital death.

Therefore, an epidemiological analysis was conducted by selecting all inpatient deaths that occurred in the hospital from 2015 to 2022. The main causes of death and specific disease composition were identified, and the SARIMA model was adopted to predict the number of new deaths. This approach allowed for a better understanding of the development trend of hospital deaths, providing insights for targeted medical service quality improvement aimed at reducing disease mortality rates.

Methods

Data collection

The mortality data were obtained from the population information registration system of the Chinese Center for Disease Control and Prevention [13, 14]. A total of 1,938 inpatient death records were collected from a tertiary general hospital in Hangzhou, covering the period from January 1, 2015, to December 31, 2022. These data were used to establish a comprehensive statistical analysis database. The main contents included in the study were sex, age, educational level, marital status, residential address, date of death and underlying cause of death. This study was approved by the Hospital Ethics Committee. The mechanisms of death were coded using the International Classification of Diseases, 10th Edition (ICD-10) [15].

Quality control

The data were obtained from the cause of death registration and reporting system of the Chinese Center for Disease Control and Prevention. The public health department of the hospital was responsible for reporting the daily Death Report Card to ensure the completeness and accuracy of the information. The review of problematic cases was conducted by a medical records team, which consists of two medical records professionals and one clinician. The hospital performs monthly quality checks on death reports. Additionally, the Hangzhou Center for Disease Control and Prevention conducted regular checks on the number of deaths in collaboration with the hospital's statistics department. They also verified the underlying cause of death by cross-checking with the medical record room. The annual rate of missing death data in hospitals is less than 2%, the rate of correct underlying cause of death codes is more than 96%, and the rate of complete identity information is more than 99%, which ensures the reliability of death data. The underreporting rate of hospital deaths is less than 2% per year, and it is concentrated in a few individuals. This under-reporting occurs randomly within the whole population and does not have any effect on the results of this study.

The study involved an analysis stratified by sex (Tables 1 and 2) to reduce the influence of confounding factors. In addition, the COVID−19 pandemic may have had an impact on the number of deaths in hospitals. The literature has reported that COVID−19 infection is linked to increased mortality in hospitalized patients with acute myocardial infarction [16]. Furthermore, studies have shown that lockdowns caused by COVID−19 are associated with a significant decrease in the number of hospitalized patients with myocardial infarction, but do not affect mortality [17]. In contrast, DZ Wang et al. reported a decrease in in-hospital mortality related to injury, poisoning, tumours, and respiratory diseases during the COVID−19 pandemic [18]. In response, an analysis was conducted to determine whether the COVID−19

Table 1 Epidemiological characteristics of deaths between 2015 and 2022 (n/%)

Characteristic	Group	Total (*n* = 1938)		Males (*n* = 1336)		Females (*n* = 602)		χ2	P
		n	%	n	%	n	%		
Year	2015	247	12.75	160	11.98	87	14.45	4.695	0.697
	2016	241	12.44	174	13.02	67	11.13		
	2017	237	12.23	168	12.57	69	11.46		
	2018	235	12.13	165	12.35	70	11.63		
	2019	223	11.51	150	11.23	73	12.13		
	2020	256	13.21	180	13.47	76	12.62		
	2021	237	12.23	159	11.90	78	12.96		
	2022	262	13.52	180	13.47	82	13.62		
Age	0–	85	4.39	45	3.37	40	6.64	46.905	< 0.001
	10–	44	2.27	26	1.95	18	2.99		
	20–	117	6.04	89	6.66	28	4.65		
	30–	145	7.48	112	8.38	33	5.48		
	40–	226	11.66	170	12.72	56	9.30		
	50–	328	16.92	246	18.41	82	13.62		
	60–	310	16.00	224	16.77	86	14.29		
	70–	282	14.55	173	12.95	109	18.11		
	80–	319	16.46	203	15.19	116	19.27		
	≥ 90	82	4.23	48	3.59	34	5.65		
Marital status	unmarried	265	13.67	185	13.85	80	13.29	61.130	< 0.001
	married	1456	75.13	1036	77.54	420	69.77		
	widowed	155	8.00	65	4.87	90	14.95		
	divorced	57	2.94	45	3.37	12	1.99		
	unknown	5	0.26	5	0.37	0	0.00		
Educational	postgraduate	4	0.21	3	0.22	1	0.17	12.310	0.031
	undergraduate	78	4.02	56	4.19	22	3.65		
	junior college	93	4.80	72	5.39	21	3.49		
	professional school	59	3.04	36	2.69	23	3.82		
	senior high school	216	11.15	165	12.35	51	8.47		
	junior high school and below	1488	76.78	1004	75.15	484	80.40		
Ethnic	Han	1914	98.76	1318	98.65	596	99.00	0.417	0.518
	minority	24	1.24	18	1.35	6	1.00		
Residential address	local area	1038	53.56	664	49.70	374	62.12	26.957	< 0.001
	other cities in the province	155	8.00	110	8.23	45	7.48		
	outside the province	745	38.44	562	42.07	183	30.40		
Season	spring	500	25.8	342	25.6	158	26.25	0.953	0.813
	summer	462	23.84	324	24.25	138	22.92		
	autumn	489	25.23	341	25.52	148	24.58		
	winter	487	25.13	329	24.63	158	26.25		

Table 2 Distribution and sex composition of patients who died from circulatory system disease, injury–poisoning, tumours and respiratory system disease (n/%)

Distribution	Total (*n* = 1938)		Males (*n* = 1336)		Females (*n* = 602)	
	n	%	*n*	%	*n*	%
circulatory system disease						
sudden cardiac death	190	30.02	144	31.86	46	25.41
acute myocardial infarction	187	29.54	138	30.53	49	27.07
cerebral haemorrhage	59	9.32	39	8.63	20	11.05
cerebral infarction	42	6.64	30	6.64	12	6.63
acute coronary syndrome	38	6.00	25	5.53	13	7.18
cardiac arrest	14	2.21	10	2.21	4	2.21
pulmonary embolism	13	2.05	7	1.55	6	3.31
injury-poisoning						
severe craniocerebral injury	265	48.45	201	50.63	64	42.67
multiple injury	74	13.53	53	13.35	21	14.00
thoracic injury	44	8.04	30	7.56	14	9.33
drowning	39	7.13	28	7.05	11	7.33
asphyxia	38	6.95	22	5.54	16	10.67
electric injury	10	1.83	9	2.27	1	0.67
severe burns	7	1.28	4	1.01	3	2.00
tumor						
lung cancer	63	22.03	41	23.16	22	20.18
liver cancer	59	20.63	42	23.73	17	15.60
lymphoma	22	7.69	14	7.91	8	7.34
pancreatic cancer	19	6.64	11	6.21	8	7.34
colorectal cancer	16	5.59	12	6.78	4	3.67
gastric cancer	14	4.90	10	5.65	4	3.67
esophageal cancer	6	2.10	6	3.39	0	0.00
respiratory system disease						
pulmonary infection	81	39.90	55	40.74	26	38.24
chronic obstructive pulmonary disease	20	9.85	14	10.37	6	8.82
hypostatic pneumonia	17	8.37	13	9.63	4	5.88
bacterial pneumonia	16	7.88	11	8.15	5	7.35
lobar pneumonia	12	5.91	9	6.67	3	4.41
interstitial pneumonia	11	5.42	7	5.19	4	5.88
respiratory failure	10	4.93	9	6.67	1	1.47

pandemic had any effect on the number of hospital deaths and is described in the discussion section.

SARIMA model

The ARIMA model is widely utilized in time series analyses for decomposing time series data into general trends, cyclical patterns, and random fluctuations [19]. The monthly time-series observations were utilized to enhance the predictive ability of the model. The SARIMA (p, d, q) (P, D, Q) s model was developed based on the ARIMA model. In the case of a nonseasonal ARIMA, the parameters include the autoregressive order (p), differencing order (d), and moving average order (q). For the SARIMA (P, D, and Q) model, the parameters include the seasonal autoregressive order (P), seasonal differencing order (D), and seasonal moving average order(Q), and s is the seasonal period length (s = 12 for 12 months). SARIMA modelling involves four key steps: stationary model fitting, parameter estimation, model diagnosis, and prediction. First, the stationarity of the SARIMA model was tested via the Dickey−Fuller test. Nonstationary time series can be differenced and seasonally differenced until they achieve stability. When the SARIMA model was constructed, the orders of (p, d, q) (P, D, Q) were determined on the basis of autocorrelation function (ACF) and partial autocorrelation function (PACF) plots. The cut−off point for the ACF plot and the slow decay in the PACF plot were used to identify the correct parameters. Additionally, the model parameters were estimated via a least-squares approach. Whether the residual sequence was white noise was determined according to the Ljung–Box Q test, and the test level was $\alpha = 0.05$. When $P > 0.05$, the residual sequence was classified as a white noise sequence. Third, diagnostic test parameters such as the mean absolute percentage error (MAPE), root mean square error (RMSE), Akaike information criterion

(AIC), and Bayesian information criterion (BIC) were employed to assess the goodness-of-fit of the constructed model and determine the optimal model [20, 21]. Finally, the SARIMA (p, d, q) (P, D, Q) s model was used for prediction, and this model was applied to predict deaths from January to December 2023. To assess the predictive ability of the model, we utilized data from 2015 to 2021 for the training set and data from 2022 for the validation set.

Statistical analysis

In this study, Excel 2016 was used to organize the death data. To address missing variables in the sample, information such as patient name, medical record number, ID number, or death date was retrieved from the hospital's medical record management system and filled in to ensure data integrity. SPSS 22.0 and R4.3.0 software (forecast, aTSA, tseries) were used to analyse the data and establish the model. The count data were analysed via frequencies and percentages, and the single factor difference was statistically inferred via χ2 tests. A two-sided $P<0.05$ indicated statistical significance.

Results

Epidemiological characteristics of deaths

In this study, 1336 patients were males (68.94%), and 602 were females (31.06%), with a sex ratio of 2.22:1. The epidemiological analysis revealed no significant difference between males and females across different years ($\chi^2=4.695$, $P=0.697$). There were variations in the age distributions of the different sexes. The percentage of males in the groups aged 20 to 29 and 30 to 39 years was significantly greater than that of females ($\chi^2=46.905$, $P<0.001$). In the widowed group, more females died than males did. However, in the married group, divorced and married males experienced a greater number of deaths than divorced and married females did ($\chi^2=61.130$, $P<0.001$). In terms of educational attainment, the proportion of males with junior college and senior high school educations was significantly greater than that of females ($\chi^2=12.310$, $P<0.05$). There was no difference in the sex distribution among the different ethnic groups ($\chi^2=0.417$, $P>0.05$). The number of deaths among males was significantly greater than that among females in out-of-province areas ($\chi^2=26.957$, $P<0.001$). There was no difference in the sex distribution across the different seasons ($\chi^2=0.953$, $P>0.05$) (Table 1).

Distribution characteristic of diseases leading to mortality from 2015 to 2022

According to the ICD−10 category statistics, the primary causes of death in the hospital were circulatory system diseases, injury, poisoning, tumours, and respiratory system diseases. These leading causes accounted for 86.12% of all deaths. The total number of deaths from circulatory diseases was 633 (452 males, 181 females), including sudden cardiac death (30.02%), acute myocardial infarction (29.54%), cerebral haemorrhage (9.32%), cerebral infarction (6.64%) and acute coronary syndrome (6.00%). Among these deaths, sudden cardiac death was the most prevalent cause of mortality in males, whereas acute myocardial infarction was the leading cause of death in females. The number of injury−poisoning deaths totalled 547, with 397 occurring in males and 150 in females. The primary causes observed were severe craniocerebral injury (48.45%), multiple injuries (13.53%), thoracic injuries (8.04%), drowning (7.13%), and asphyxia (6.95%). The distributions of major diseases were consistent between males and females. There were 286 deaths from malignant tumours (177 males, 109 females), which included lung cancer (22.03%), liver cancer (20.63%), lymphoma (7.69%), pancreatic cancer (6.64%) and colorectal cancer (5.59%). Liver cancer was the leading cause of death in males, whereas lung cancer was the leading cause of death in females. In total, there were 203 deaths from respiratory diseases, with 135 occurring in males and 68 occurring in females. The main causes of these deaths were pulmonary infection (39.90%), chronic obstructive pulmonary disease (9.85%), hypostatic pneumonia (8.37%), bacterial pneumonia (7.88%), and lobar pneumonia (5.91%). Among these causes, pulmonary infections was the main cause of respiratory system−related death in both males and females (Table 2).

SARIMA model

Monthly death cases were utilized to establish a SARIMA model covering the period from January 2015 to December 2021. The data were organized, and a sequence chart was created. The average annual number of deaths was 242. Figure 1 shows the monthly deaths from January 2015 to December 2021, revealing no significant trend of change. The number of deaths during this period was analysed via deterministic factors, and the original data were confirmed to be a nonstationary sequence via the Dickey−Fuller test ($P>0.05$). Moreover, the data exhibited seasonal trends. The Ljung–Box Q test demonstrated that the sequence was not a purely random sequence but rather a nonwhite noise sequence ($P<0.05$) (Fig. 2). Therefore, a SARIMA model could be established.

To achieve stationarity in the series, first-order differencing adjustments were applied both seasonally and nonseasonally (Fig. 3). Figure 4 shows that the ACF and PACF of the new data tended to be stationary after the first-order seasonal difference adjustment was applied. This indicates that the values of d and D in the SARIMA $(p, 1, q) \times (P, 1, Q)_{12}$ model were 1 and 1, respectively. For the fitting of the respective models, a combined spectrum of parameters was compared for the SARIMA (p, 1,

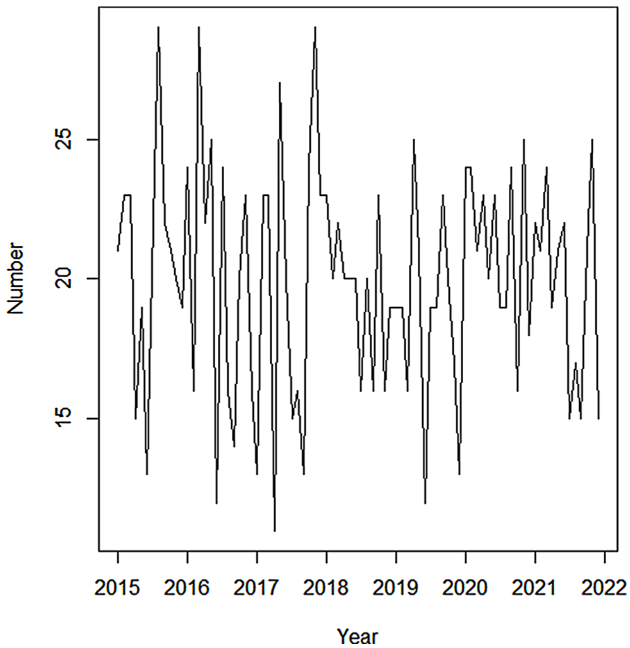

Fig. 1 Sequence chart of hospital deaths from January 2015 to September 2021

q) × (P, 1, Q)$_{12}$ model, and an optimal model was selected according to the criteria of the minimum AIC, BIC and MAPE. The final model for deaths was the SARIMA (2, 1, 1) (1, 1, 1)$_{12}$ model, which showed the best goodness of fit, with an AIC of 380.23, a BIC of 392.79, an AICc of 381.81, a MAPE of 14.99%, and an RMSE of 3.83. The residuals in the fitted SARIMA models were found to be pure random sequences, as confirmed by the Ljung–Box Q test (P=0.956).

The SARIMA (2, 1, 1) (1, 1, 1)$_{12}$ model was employed to analyse the number of deaths in the hospital from January 2015 to December 2021. The results indicated that the SARIMA (2, 1, 1) (1, 1, 1)$_{12}$ model effectively predicted the number of deaths within a short time (Fig. 5). The model was used to predict the number of deaths in 2022, and the actual value was employed to assess the model's accuracy. Table 3 indicates that the 95% confidence interval for nearly all the predicted values encompassed the actual values, with a relative error of 8.02%, demonstrating a strong fit for the model. Finally, the SARIMA (2, 1, 1) (1, 1, 1)$_{12}$ model was used to predict the number of deaths among hospital patients from January to December 2023. The model predicted a total of 241 deaths during this period.

Discussion

The number of inpatient deaths in the hospital from 2015 to 2022 showed minimal overall change, with the figure remaining steady at approximately 242 cases. The trend of inpatient deaths in a top-three hospital in Urumqi [22] was similar to the change trend but differed from the trend observed in a top-three hospital in Shenzhen. In the latter case, there was a consistent downward trend in the number of inpatient deaths from 2014 to 2019 [23]. This study revealed that the establishment of a critical

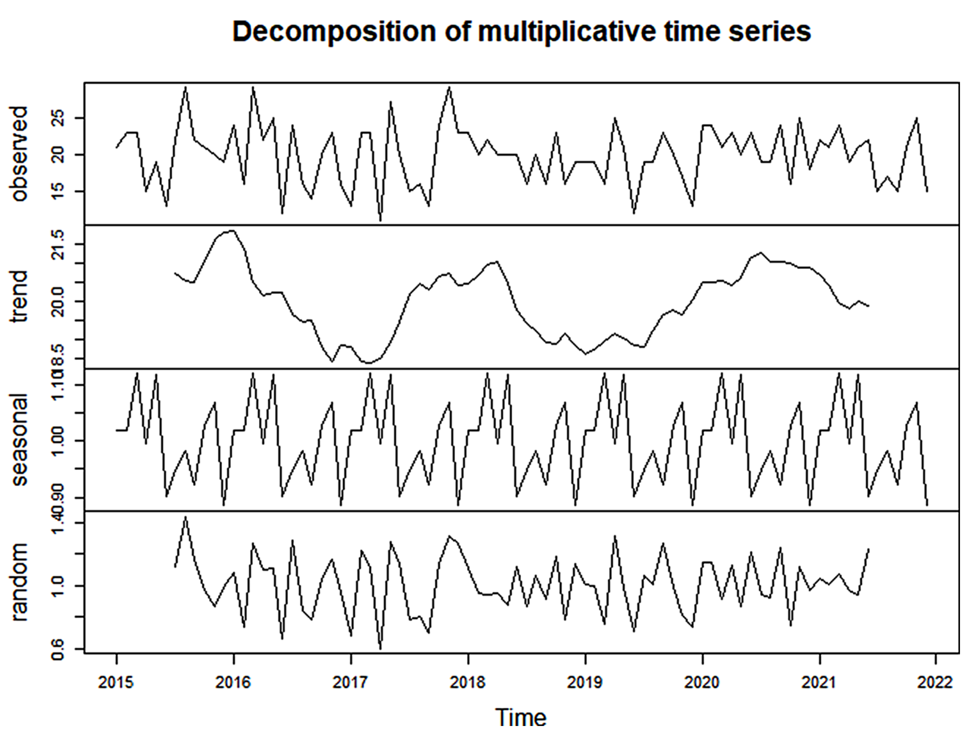

Fig. 2 Deterministic analysis chart of hospital deaths from January 2015 to September 2021

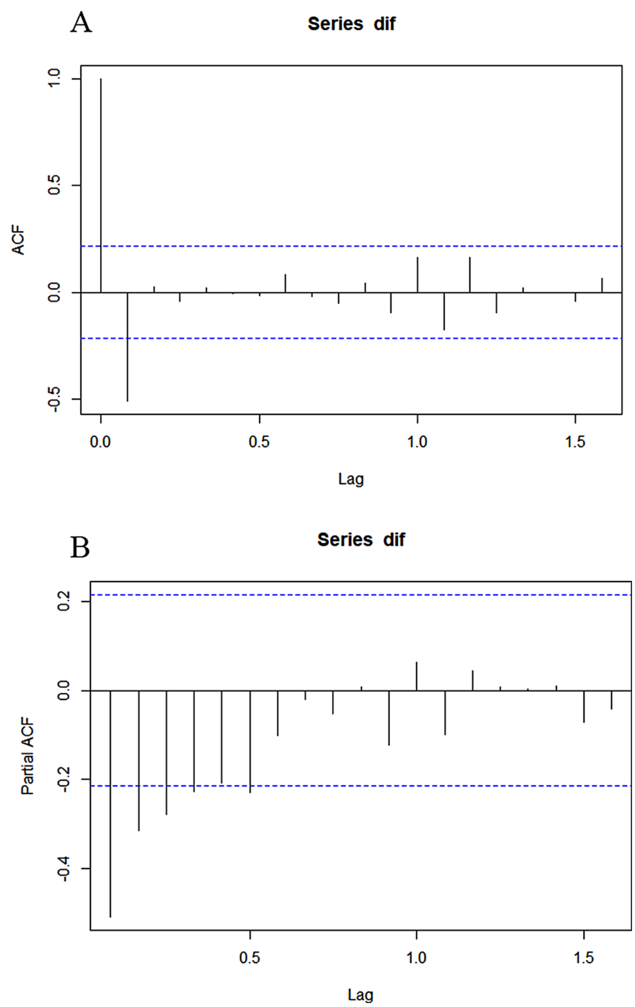

Fig. 3 ACF and PACF charts of the number of hospital deaths from January 2015 to September 2021

care centre in third-level hospitals in Shenzhen improved the ability to diagnose, treat and rescue critical care diseases. This has led to a gradual reduction in the number of inpatient deaths. Reducing mortality and improving

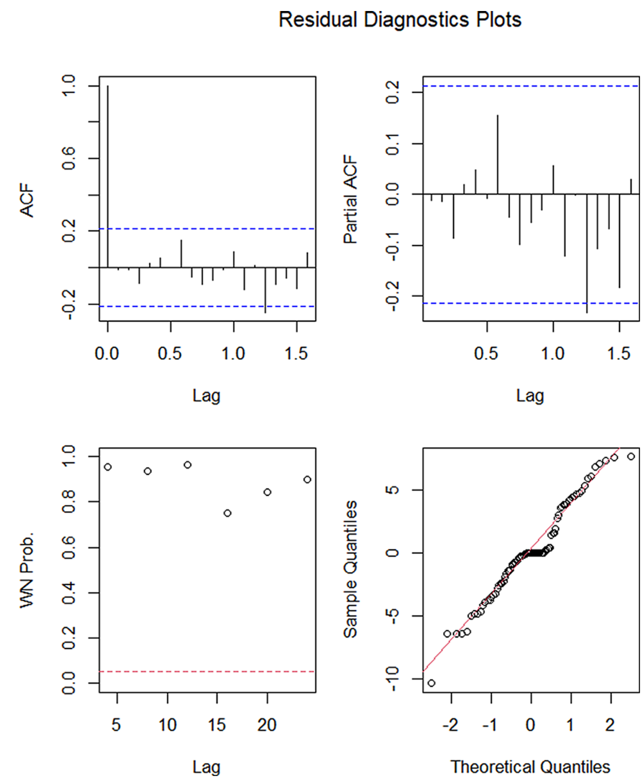

Fig. 4 Significance test diagram of the model fit

the health of hospitalized patients are common objectives of hospitals. Notably, the most severe COVID-19 outbreaks occurred in 2020, posing significant challenges for health care facilities worldwide. Compared with that in previous years, the number of deaths from cardiovascular and cerebrovascular diseases has increased. However, there has been a decline in injury–poisoning deaths. The number of tumour deaths has not fluctuated significantly. Respiratory deaths have been increasing. COVID–19 required the use of some medical resources, seriously affecting the hospitalization of patients with critical diseases such as acute myocardial infarction. As a result,

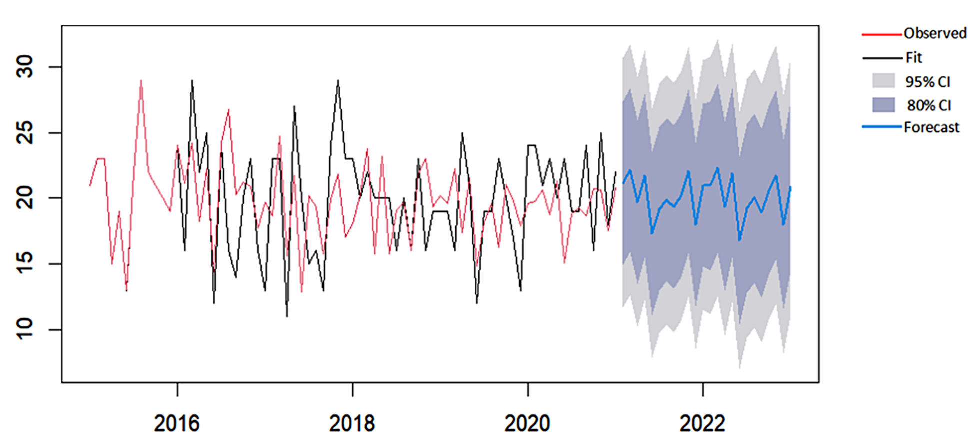

Fig. 5 Fitting and prediction of the SARIMA (2, 1, 1) $(1, 1, 1)_{12}$ model

Table 3 Comparison between actual values and predicted values of the number of deaths from the SARIMA model in 2022

Month	Predicted value	Actual value	Absolute error	Relative error (%)	95%CI
1	21	16	-5	-31.25	11.54–30.48
2	21	17	-4	-23.53	11.13–30.79
3	22	17	-5	-29.41	12.49–32.15
4	19	18	-1	-5.56	9.56–29.20
5	22	19	-3	-15.79	12.10-31.74
6	17	27	10	37.04	6.98–26.62
7	19	18	-1	-5.56	9.41–29.06
8	20	29	9	31.03	10.17–29.82
9	19	15	-4	-26.67	9.07–28.71
10	21	26	5	19.23	10.75–30.40
11	22	21	-1	-4.76	11.94–31.57
12	18	39	21	53.85	8.20-27.83

the level of treatment decreased after the pandemic compared with that before pandemic [24]. The decrease in the rate of mortality due to motor vehicle traffic accidents and accidental falls may be attributed to a decrease in exposure to related risk factors due to vehicle restrictions, and production shutdowns during the pandemic period [18]. The spread of COVID−19 during an outbreak increased the risk of respiratory illness and death [25].

Male patients had a higher mortality rate than did the female patients in the hospital (2.22:1), which aligns with the findings of Salaj D and Campbell JE et al. [26, 27]. Males are more likely than females to have unhealthy lifestyles, such as smoking, drinking, staying up late, engaging in occupations in high−risk industries and engaging in high-intensity physical activity. It is important to pay more attention to men's health problems to reduce male mortality. There were differences in age, marital status, education level, residence and other aspects of death between the sexes ($P<0.05$), which aligns with the findings of Wang Z's study [22]. The distribution of deaths spanned all age groups. Among these deaths, 6.66% occurred in patients aged 0 to 19 years, primarily because of congenital heart disease and accidental injury. The group aged 20 to 59 years accounted for 42.10% of the deaths, with causes including sudden death, acute myocardial infarction, and traffic accidents. Those aged 60 years or older accounted for the largest proportion at 51.24%, with causes such as sudden cardiac death, cerebral infarction, cancer, and pulmonary infection. Therefore, hospitals should focus on strengthening the ability to rescue patients with acute and critical illnesses and establish and improve prevention and treatment systems for senile diseases.

On the basis of the results of the statistical analysis according to the ICD−10 classification, the primary causes of death in the hospital were circulatory system diseases, injury, poisoning, tumours, and respiratory diseases. These findings are consistent with those of previous domestic studies and reflect the current trend of disease development in China [28, 29]. Cardiovascular disease is the leading cause of death worldwide [30]. This study revealed that the main causes of death from circulatory diseases were sudden cardiac death, acute myocardial infarction and cerebral haemorrhage. Huang S et al. [31] reported that the annual incidence of sudden cardiac death in China is approximately 41.84 deaths per 100,000 people, making China the country with the highest incidence of sudden cardiac death worldwide. In addition, injury−poisoning contributes significantly to the disease burden on society. Mulima G et al. [32] reported that over five million individuals succumb to injuries annually, constituting 9% of all fatalities worldwide. The study revealed that traffic accidents accounted for approximately 40% of incidents, falls from heights accounted for approximately 30%, and drownings accounted for approximately 10%. This trend is closely associated with the significant increase in the number of motor vehicles, rapid growth in the construction industry, illegal driving practices, and a lack of public awareness regarding traffic safety. The global burden of malignant tumours is substantial and continues to increase. In our hospital, lung cancer accounted for a significant proportion (22.03%) of deaths from malignant tumours, which aligns with the findings of Zhang JY [33] and Ha L et al. [34]. The occurrence of liver cancer, pancreatic cancer, colorectal cancer, and other digestive system tumours may be associated with an unbalanced diet, excessive drinking, overeating, and a lack of exercise. Deaths from respiratory diseases included deaths due to pulmonary infections, chronic obstructive pulmonary disease, and hypostatic pneumonia. These findings are consistent with the research of Halpin DMG et al. [35–37]. Studies have shown that as urban air pollution intensifies and population density increases, there is a greater likelihood of the spread of viruses and bacteria, leading to an increased incidence of respiratory diseases [38].

The SARIMA model not only considers seasonal periodicity but also captures nonseasonal components on the basis of sequential changes within the seasonal cycle. This feature enhances the practicality and accuracy of the model, making it suitable for academic research and analysis [39, 40]. Yang WJ et al. [41] used the SARIMA (2, 2, 2) $(0, 1, 1)_{12}$ model to predict the number of inpatients in the top three hospitals in Zhejiang Province, and the prediction effect was good. Liu JC et al. [42] modelled and predicted the number of inpatients with acute mountain sickness (AMS) via the ARIMA seasonal product model and ultimately determined that the ARIMA (1, 1, 1) $(1, 0, 1)_{12}$ model was the optimal model. In this study, a preliminary model was established on the basis of the ACF and PACF, considering both seasonal and nonseasonal factors. After many attempts, the ARIMA (2, 1, 1) $(1, 1, 1)_{12}$ model was identified as the optimal model. The relative error between the predicted value and the actual value in 2022 was 9.54%, indicating the strong forecasting ability of the model. Compared with previous years, the SARIMA (2, 1, 1) $(1, 1, 1)_{12}$ model predicted 241 deaths in 2023 with little fluctuation. The analysis of this time series model not only helps analyse the impact of changes in the number of deaths in hospitals but also provides a reference for how to reduce the case fatality rate in hospitals after the end of the pandemic.

According to the aforementioned findings, the following recommendations for hospital management are proposed. First, given the significance of response time in cases of cardiovascular and cerebrovascular diseases, it is crucial to conduct regular first-aid drills and enhance first-aid proficiency. Second, there is a need to increase the capacity for multidisciplinary collaboration in treatment and establish a streamlined process for trauma care. Additionally, efforts should be made to increase cancer screening, early detection, and treatment for high-risk populations to reduce disease mortality rates. It is important to acknowledge certain limitations of this study. First, the in-hospital death data were incomplete because some patients and their family members discontinued treatment, leading to information bias. Second, our study only analysed hospital death data from 2015 to 2022, which may reduce the accuracy of the SARIMA model's predictions as the time horizon extended. Finally, the data was obtained from a tertiary general hospital in Hangzhou, which can reflect only the death situation within the hospital's jurisdiction. Analysing the death case data solely from the perspective of one hospital is insufficient to fully represent the overall mortality situation in the entire Hangzhou area. Therefore, the representativeness of this study is limited. In the future, more comprehensive death data will be collected to increase the specificity and depth of the study. Furthermore, future research will explore the use of the SARIMA model in developing a predictive model for the incidence of specific diseases. Comparing the SARIMA model with other models, such as the seasonal autoregressive fractionally integrated moving average (SARFIMA) model, the SARIMA-ETS-SVR hybrid model, and the Holt-Winters model, can help in selecting a model with higher predictive efficiency [21, 43, 44].

Conclusions

An analysis of the causes of death in the hospital over an 8-year period revealed that sudden cardiac death, acute myocardial infarction, severe craniocerebral injury, lung cancer, and pulmonary infection were the primary causes of fatalities observed in the hospital. The SARIMA (2, 1, 1) $(1, 1, 1)_{12}$ model is suitable for predicting the number of hospital deaths and can serve as a basis for the rational allocation of medical resources in hospitals.

Abbreviations

ARIMA	Autoregressive integrated moving average
ICD-10	International Classification of Diseases 10th Edition
ACF	Autocorrelation function
PACF	Partial autocorrelation function
MAPE	Mean absolute percentage error
RMSE	Root mean square error
AIC	Akaike information criterion
BIC	Bayesian information criterion
AMS	Acute mountain sickness
SARFIMA	Seasonal autoregressive fractionally integrated moving average

Acknowledgements

We sincerely thank all the study participants. The authors declare that they have no conflicting interests. No conflict ofinterest exist in the submission of this manuscript, and manuscript is approved by authors for publication. All the authors lists have approved the manuscript that is enclosed.

Author contributions

Jingyuan Dai: study design, data interpretation, manuscript preparation and revision. Yun Xiao: data analysis review. Qionglian Sheng: review of relevant literature. Jing Zhou: data acquisition and analysis. Zhe Zhang: critical revision of the manuscript for important intellectual content. Fenglong Zhu: manuscript revision and interpretation of the results. All the authors contributed to the writing of the manuscript and approved the manuscript for submission.

Funding

Not applicable.

Data availability

Data is provided within the manuscript.

Declarations

Ethics approval and consent to participate

This study protocol was reviewed and approved by the Ethics Committee of Linping Campus, Second Affiliated Hospital, Zhejiang University School of Medicine (Clinical Trial Number: 2023/017). Written informed consent was obtained from all participants.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Department of Case Statistics, Second Affiliated Hospital, Zhejiang University School of Medicine, Linping Campus, Hangzhou 311199, China

[2]Department of Quality Management, Second Affiliated Hospital, Zhejiang University School of Medicine, Linping Campus, Hangzhou 311199, China
[3]Department of Medical Laboratory, Second Affiliated Hospital, Zhejiang University School of Medicine, Jiande Campus, Hangzhou 311600, China

Received: 28 December 2023 / Accepted: 10 September 2024
Published online: 19 September 2024

References

1. Liao R, Hu L, Liao Q, et al. Analysis of death causes of residents in poverty-stricken areas in 2020: take Liangshan Yi Autonomous Prefecture in China as an example. BMC Public Health. 2022;22:89.
2. Gong L. Analysis of 7831 deaths in a Tertiary Hospital. Chin J Med. 2023;24(09):83–6.
3. Li Y. Analysis of death cases of inpatients in a Tertiary Hospital in Zhejiang Province from 2017 to 2021. Med Inform. 2023;36(15):88–91.
4. Du J, Sun Z. Epidemiological analysis of 3682 deaths from 2005 to 2019 in a third-grade hospital in Beijing. Chin Med J. 2022;23(05):67–71.
5. Chen FF, Wen JZ, Xu SM. Analysis of 6833 deaths in a tertiary A hospital from 2003 to 2018. Chin Hosp Stat. 2020;27(02):119–23.
6. Chen Y. Epidemiological characteristics of in-patient deaths in a top three hospital in Xi'an from 2015 to 2019. Clin Med Res Pract. 2022;7(01):10–3.
7. Li Y, Ning Y, Shen B, et al. Temporal trends in prevalence and mortality for chronic kidney disease in China from 1990 to 2019: an analysis of the global burden of Disease Study 2019. Clin Kidney J. 2023;16:312–21.
8. Wang M, Pan J, Li X, et al. ARIMA and ARIMA-ERNN models for prediction of pertussis incidence in mainland China from 2004 to 2021. BMC Public Health. 2022;22:1447.
9. Zheng A, Fang Q, Zhu Y, et al. An application of ARIMA model for predicting total health expenditure in China from 1978–2022. J Glob Health. 2020;10:010803.
10. James A, Tripathi V. Time series data analysis and ARIMA modeling to forecast the short-term trajectory of the acceleration of fatalities in Brazil caused by the corona virus (COVID-19). Peer J. 2021;9:e11748.
11. Qi F, Xu ZS, Zhang H. Predicting the mortality of smoking attributable to cancer in Qingdao, China: a time-series analysis. PLoS ONE. 2021;16:e0245769.
12. Guo J, Zhang L, Guo R. Relative humidity prediction with covariates and error correction based on SARIMA-EG-ECM model. Model Earth Syst Environ. 2023; 1–13.
13. Sha F, Chang Q, Law YW. Suicide rates in China, 2004–2014: comparing data from two sample-based mortality surveillance systems. BMC Public Health. 2018;18:239.
14. Liu W, Liu J, Song Y. Mortality of lymphoma and myeloma in China, 2004–2017: an observational study. J Hematol Oncol. 2019;12:22.
15. Figueroa CA, Linhart CL, Dearie C. Effects of inappropriate cause-of-death certification on mortality from cardiovascular disease and diabetes mellitus in Tonga. BMC Public Health. 2023;23:2381.
16. Baral N, Abusnina W, Balmuri S. COVID-19 positive status is sssociated with increased in-hospital mortality in patients with acute myocardial infarction: a systematic review and meta-analysis. J Community Hosp Intern Med Perspect. 2022;12:17–24.
17. Baumhardt M, Dreyhaupt J, Winsauer C. The effect of the lockdown on patients with myocardial infarction during the COVID-19 pandemic-a systematic review and meta-analysis. Dtsch Arztebl Int. 2021;118:447–53.
18. Wang DZ, Zhang S, Zhang H. New residents in Tianjin crown pneumonia strictly control strategy because of the influence of death. China's Chronic Disease Prev Control. 2021;29(11):801–7.
19. Costa EM, Magalhães RES. The Brazilian national oral health policy and oral cancer mortality trends: an autoregressive integrated moving average (ARIMA) model. PLoS ONE. 2023;18:e0291609.
20. Ab RMA, Ahmad ZR, Wan MWR. Forecasting new tuberculosis cases in Malaysia: a Time-Series Study using the Autoregressive Integrated moving average (ARIMA) model. Cureus. 2023;15:e44676.
21. Wang YB, Qing SY, Liang ZY. Time series analysis-based seasonal autoregressive fractionally integrated moving average to estimate hepatitis B and C epidemics in China. World J Gastroenterol. 2023;29:5716–27.
22. Wang Z. Analysis and prediction of death cases in a top three hospital in Urumqi area. Xinjiang medical university; 2022.
23. Liang JF, Qiu JZ, Li M. Analysis of inpatient death cases in a top three hospital in Shenzhen from 2014 to 2019. Chin J Health Stat. 2019;38(03):425–7.
24. De Rosa S, Spaccarotella C, Basso C, et al. Reduction of hospitalizations for myocardial infarction in Italy in the COVID-19era. EurHeart J. 2020;41(22):2083–8.
25. Giesing W, Soney H, Wang L. Outcomes of hospitalised COVID-19 patients arriving with hypoxic respiratory failure. Lung and Circulation: Heart; 2024.
26. Salaj D, Schultz T, Strang P. Nursing home residents with dementia at end of life: emergency department visits, hospitalizations, and acute hospital deaths. J Palliat Med. 2024;27(1):24–30.
27. Campbell JE, Sambo AB, Hunsucker LA. Rural cancer disparities from Oklahoma cancer and vital records registries 2016–2020. Cancer Epidemiol. 2023;88:102512.
28. Varela DV, Martins MRO, Furtado A. Spatio-temporal evolution of mortality in Cape Verde: 1995–2018. Plos glob Public Health. 2023;3:e0000753.
29. Wang DZ, Zhang H, Zhang S. Study on increase of average life expectancy of residents in Tianjin from 1999 to 2018. Chin J Epidemiol. 2021;42:814–22.
30. Litviňuková M, Talavera-López C, Maatz H. Cells of the adult human heart. Nature. 2020;588:466–72.
31. Huang S, Zhang J, Wan H. Plasma extracellular vesicles microRNA-208b-3p and microRNA-143-3p as novel biomarkers for sudden cardiac death prediction in acute coronary syndrome. Mol Omics. 2023;19(3):262–73.
32. Mulima G, Purcell LN, Maine R. Epidemiology of prehospital trauma deaths in Malawi: a retrospective cohort study. Afr J Emerg Med. 2021;11:258–62.
33. Zhang JY, Wang YN, Yuan BW. Identifying key transcription factors and immune infiltration in non-small-cell lung cancer using weighted correlation network and Cox regression analyses. Front Oncol. 2023;13:1112020.
34. Ha L, Tran A, Bui L. Proportion and number of cancer cases and deaths attributable to behavioral risk factors in Vietnam. Int J Cancer. 2023;153:524–38.
35. Roig-Marín N, Roig-Rico P. Ground-glass opacity on emergency department chest X-ray: a risk factor for in-hospital mortality and organ failure in elderly admitted for COVID-19. Postgrad Med. 2023;3:265–72.
36. Halpin DMG, Martinez FJ. Pharmacotherapy and mortality in chronic obstructive pulmonary disease. Am J Respir Crit Care Med. 2022;206(10):1201–7.
37. Roig-Marín N, Roig-Rico P. The deadliest lung lobe in COVID-19: a retrospective cohort study of elderly patients hospitalized for COVID-19. Postgrad Med. 2022;5:533–9.
38. Zhang JW, Lim Y, So R. Long-term exposure to air pollution and risk of SARS-CoV-2 infection and COVID-19 hospitalisation or death: Danish nationwide cohort study. Eur Respir J. 2023; 62.
39. Agyemang EF, Mensah JA, Ocran E, et al. Time series based road traffic accidents forecasting via SARIMA and Facebook Prophet model with potential changepoints. Heliyon. 2023;9:e22544.
40. Li CL, Cui ZZ, Wei D, et al. Trends and patterns of Antibiotic prescriptions in Primary Care Institutions in Southwest China, 2017–2022. Infect Drug Resist. 2023;16:5833–54.
41. Yang WJ, Su A, Ding LP. Application of exponential smoothing method and SARIMA model in predicting the number of admissions in a third-class hospital in Zhejiang Province. BMC Public Health. 2023;23:2309.
42. Liu JC, Yu FF, Song H. Application of SARIMA model in forecasting and analyzing inpatient cases of acute mountain sickness. BMC Public Health. 2023;23:56.
43. Zhao D, Zhang R. A new hybrid model SARIMA-ETS-SVR for seasonal influenza incidence prediction in mainland China. J Infect Dev Ctries. 2023;17:1581–90.
44. Xian XB, Wang L, Wu XH. Comparison of SARIMA model, Holt-winters model and ETS model in predicting the incidence of foodborne disease. BMC Infect Dis. 2023;23:803.

Publisher's note