

RESEARCH ARTICLE

Trends, structural changes, and assessment of time series models for forecasting hospital discharge due to death at a Mexican tertiary care hospital

Edel Rafael Rodea-Montero^{1*}, Rodolfo Guardado-Mendoza¹, Brenda Jesús Rodríguez-Alcántar¹, Jesús Rubén Rodríguez-Núñez², Carlos Alberto Núñez-Colín², Lina Sofía Palacio-Mejía³

1 Department of Research, Hospital Regional de Alta Especialidad del Bajío, León, Guanajuato, México, **2** Programa de Biotecnología, Universidad de Guanajuato, Celaya, Guanajuato, México, **3** CONACYT-Instituto Nacional de Salud Pública, Cuernavaca, Morelos, México

* edel.rodea@hraeb.gob.mx



OPEN ACCESS

Citation: Rodea-Montero ER, Guardado-Mendoza R, Rodríguez-Alcántar BJ, Rodríguez-Núñez JR, Núñez-Colín CA, Palacio-Mejía LS (2021) Trends, structural changes, and assessment of time series models for forecasting hospital discharge due to death at a Mexican tertiary care hospital. PLoS ONE 16(3): e0248277. <https://doi.org/10.1371/journal.pone.0248277>

Editor: Qiang Zeng, South China University of Technology, CHINA

Received: December 15, 2020

Accepted: February 23, 2021

Published: March 8, 2021

Copyright: © 2021 Rodea-Montero et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All data underlying the findings are available in the Mexican General Directorate of Health Information website at: http://www.dgis.salud.gob.mx/contenidos/basesdedatos/da_egresoshosp_gobmx.html.

Funding: The author(s) received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

Abstract

Background

Data on hospital discharges can be used as a valuable instrument for hospital planning and management. The quantification of deaths can be considered a measure of the effectiveness of hospital intervention, and a high percentage of hospital discharges due to death can be associated with deficiencies in the quality of hospital care.

Objective

To determine the overall percentage of hospital discharges due to death in a Mexican tertiary care hospital from its opening, to describe the characteristics of the time series generated from the monthly percentage of hospital discharges due to death and to make and evaluate predictions.

Methods

This was a retrospective study involving the medical records of 81,083 patients who were discharged from a tertiary care hospital from April 2007 to December 2019 (first 153 months of operation). The records of the first 129 months (April 2007 to December 2017) were used for the analysis and construction of the models (training dataset). In addition, the records of the last 24 months (January 2018 to December 2019) were used to evaluate the predictions made (test dataset). Structural change was identified (Chow test), ARIMA models were adjusted, predictions were estimated with and without considering the structural change, and predictions were evaluated using error indices (MAE, RMSE, MAPE, and MASE).

Results

The total percentage of discharges due to death was 3.41%. A structural change was observed in the time series (March 2009, $p > 0.001$), and $ARIMA(0,0,0)(1,1,2)_{12}$ with drift

models were adjusted with and without consideration of the structural change. The error metrics favored the model that did not consider the structural change (MAE = 0.63, RMSE = 0.81, MAPE = 25.89%, and MASE = 0.65).

Conclusion

Our study suggests that the ARIMA models are an adequate tool for future monitoring of the monthly percentage of hospital discharges due to death, allowing us to detect observations that depart from the described trend and identify future structural changes.

Introduction

Data on hospital discharges due to death can be used as a valuable instrument for hospital planning and management [1]. In Mexico, the General Directorate of Health Information (*Dirección General de Información en Salud*—DGIS; Spanish acronym) is the operational body of the Ministry of Health (*Secretaría de Salud*—SSA; Spanish acronym) that is responsible for generating statistics on health and has various information subsystems, including the Automated Hospital Discharge System (*Sistema Automatizado de Egresos Hospitalarios*—SAEH; Spanish acronym) [2]. The number of national hospital discharges recorded by Mexican SSA hospitals as primary sources in the SAEH is estimated to be 3 million cases per year, and of these, approximately 2% are deaths [3]. There are descriptive reports on national and regional hospital discharges [2,4], but there are few inferential studies on trends, the identification of associated variables, or predictions about such hospital discharges.

The Bajío Regional Hospital of High Specialty (*Hospital Regional de Alta Especialidad del Bajío*—HRAEB) has provided clinical, diagnostic, and tertiary treatment services since April 2007. It has 184 census beds, does not have an emergency department, and discharges patients with complex pathologies. Its source documents for recording hospital discharges are the clinical file and death certificate. The corresponding data from on these documents are captured in the computer platform of the SAEH by the Medical Statistics area personnel of the HRAEB according to the guidelines of the DGIS [5].

Since 1986, the Health Care Financing Administration has incorporated the hospital mortality rate as a qualitative comparator of American hospitals, which has encouraged the use of hospital care outcome indicators across the world [6]. Hospital mortality is one of the most frequently used indicators of quality of care [7], and the quantification of deaths can be considered a measure of the effectiveness of hospital intervention, although it should not be forgotten that this indicator is influenced by other factors, such as the pathology being treated, population structure and accessibility to the hospital unit [8]. A high percentage of hospital discharges due to death can be associated with deficiencies in the quality of hospital care [9].

Since in this case the percentage of discharges due to death it is a change of percentage over time, it can be defined as a time series. To model this mortality series, linear models are typically constructed [10]. Some studies have even analyzed the trends of hospital discharges for a specific cause [11], and other authors have estimated predictions [12]. Usually Autoregressive integrated moving average (ARIMA) models are used for analyzing and forecasting time series data and are also capable of modelling a wide range of seasonal data [13–15]. The Box-Jenkins (ARIMA) econometric modelling is a forecasting technique that completely ignores independent variables in making forecast. It takes into account historical data and decomposes it into Autoregressive (AR) process, where there is a memory of past events; an Integrated (I) process, which accounts for stabilizing or making the data stationary, making it easier to forecast; and a

Moving Average (MA) of the forecast errors, such that the longer the historical data, the more accurate the forecasts will be, as it learns over time [16,17].

For the identification of the ARIMA model, two goodness-of-fit statistics that are most commonly used for the model selection are; Akaike Information Criterion (AIC) and Schwarz Bayesian Information Criterion (BIC). The AIC and BIC are determined based on a likelihood function [15]. When comparing two models, the one with the lower AIC or BIC is generally better [18]. Low AIC or BIC values suggest that a model nicely straddle the requirements of goodness-of-fit and parsimony [19].

In addition, regarding hospital deaths, there is scientific evidence that certain events, including enabling hospital admissions on weekends [20,21], varying the number and educational level of nursing staff as well as the ratio of nurses per census bed [22–25], increasing the number of hospitalized patients [26,27] and increasing the volume of surgical patients [28,29], can increase the risk of hospital mortality and therefore raise the percentage of hospital discharges due to death.

The aim of this study was to determine the overall percentage of hospital discharges due to death in a Mexican tertiary care hospital from its opening in April 2007 to December 2019, to describe the characteristics of the time series generated from the monthly percentage of hospital discharges due to death from April 2007 to December 2017 (first 129 months of operation) and to make and evaluate predictions of the monthly percentage of hospital discharges due to death from January 2018 to December 2019 (24 months later).

Materials and methods

Patients

This retrospective study of Hispanic-Mexican patients included all the records of patients ($n = 81,083$) who were discharged from a tertiary care hospital (HRAEB) located in the city of León in the state of Guanajuato, Mexico, from its opening in April 2007 to December 2019 (153 months of operation). The dataset is secondary and was obtained from the national government records of the SAEH subsystem operated by DGIS [3]. This is a subsystem that compiles national hospital discharge information from Mexican hospitals as primary sources. The dataset was divided into two parts. The records of the first 129 months (April 2007 to December 2017) were used for the analysis and construction of the models (training dataset). In addition, the records of the last 24 months (January 2018 to December 2019) were used to evaluate the predictions made (test dataset).

Ethical considerations

The protocol of this study was reviewed and approved by the Institutional Research and Ethics Committees of the HRAEB and the National Institute of Public Health (*Instituto Nacional de Salud Pública*) of Mexico (approval numbers: CI/HRAEB/2019/046 and PT 211, respectively).

Statistical analysis

All data were analyzed with the statistical software R [30]. Initially, a descriptive analysis was implemented, and the time series associated with the monthly percentage of hospital discharges due to death was plotted. For the time series, the trend was modeled using a simple linear regression model using least squares (SLRMLS) without segmentation [10]. The assumptions of normality, homoscedasticity, and independence of the residuals were verified using the Kolmogorov-Smirnov, White, and Durbin-Watson tests, respectively. Next, the instances where possible structural changes could occur (sudden changes in the trend of the

time series) were identified using the Chow test [31] with the statistical R package “struc-change”, which is designed to test structural changes in linear regression models [32]. Next, additive decomposition of the series into its components was performed as follows: trend, seasonal variations (seasonality), and irregular fluctuations (random) [10]. ARIMA models were used. The name of this procedure comes from its three components: autoregression (AR), integration (I), and moving average (MA). The ARIMA(p, d, q) model is fitted to the observed data, where p, d and q represent the order of the three respective components. Regarding the integrated components of the ARIMA model, if the stochastic process associated with a time series has a unit root, we can conclude that it is a nonstationary time series. Therefore, the Dickey-Fuller test was implemented for each of the y_t series to determine if it was stationary. If not, we proceeded to differentiate each series (generating a new series $z_t = y_t - y_{t-1}$ with $t = 2, 3, \dots, n$) as many times as necessary (d) to obtain a stationary time series and thus conclude that the original series y_t was integrated of order (d). In addition, autocorrelation (ACF) and partial autocorrelation (PACF) plots were constructed for each series and for those series that resulted from differentiation [33] to empirically confirm whether the series could be considered stationary and identify the possible number of lags for each of the remaining components of the ARIMA model.

In order to be able to predict the monthly percentage of discharges due to death, with and without considering the structural change identified in the series, for the forecasting, the seasonal ARIMA models were adjusted. The seasonal ARIMA(p, d, q)(P, D, Q)_s process is given by $\Phi(B^s)\phi(B)(1-B^s)^D(1-B)^d y_t = c + \Theta(B^s)\theta(B)\epsilon_t$ where $\Phi(z)$ and $\Theta(z)$ are polynomials of orders P and Q respectively, each containing no roots inside the unit circle. If $c \neq 0$, there is an implied polynomial of order $d + D$ in the forecast function [34]. The selection of the proper seasonal ARIMA model in each case was done by the `auto.arima` function from the R package “forecast” [35]. This function returns the best ARIMA model according to the AIC or BIC value [36]. The function conducts a search over a range of possible models within the order constraints and picks the most suitable one [37]. The `auto.arima` function is based mainly on the Hyndman–Khandakar algorithm, which combines unit element tests, minimization of corrected Akaike tests (AICc), and the maximum likelihood method to obtain the model most appropriate for the data. The criteria of goodness-of-fit, based on the information criterion, were taken into account [38]: $AIC = -2\ln(L) + 2(p+q+P+Q+k)$ and $BIC = -2\ln(L) + (p+q+P+Q+k)\ln(n)$ where $k = 1$ if $c \neq 0$ and 0 otherwise, and L is the maximized likelihood of the model fitted to the differenced data $(1-B^s)^D(1-B)^d y_t$. After selecting the models for each series, it was verified that the residuals did not have a dependence structure and that they followed a white noise process.

Finally, predictions were made (point forecasts and prediction intervals) of the series of the percentage of discharges due to death, with and without considering the structural change identified in the series using the function `forecast` from the R package “forecast” [35]. These predictions were evaluated using four error indices: two scale-dependent error indices: mean absolute error (MAE) and root mean square error (RMSE); one percentage error index: mean absolute percentage error (MAPE); and one scale-free error index: mean absolute scaled error (MASE) [39–42]. In all tests, the significance level of $\alpha = 0.05$ was used.

Results

A total of 81,083 hospital discharges, of which 2,767 were deaths, were generated in the HRAEB from April 2007 to December 2019 and included in the final analysis of this study. [Table 1](#) describes the total hospital discharges of the HRAEB per year, and a total percentage of discharges due to death of 3.41% was observed. In 2009, the highest annual percentage of

Table 1. Hospital discharges in the HRAEB grouped by year.

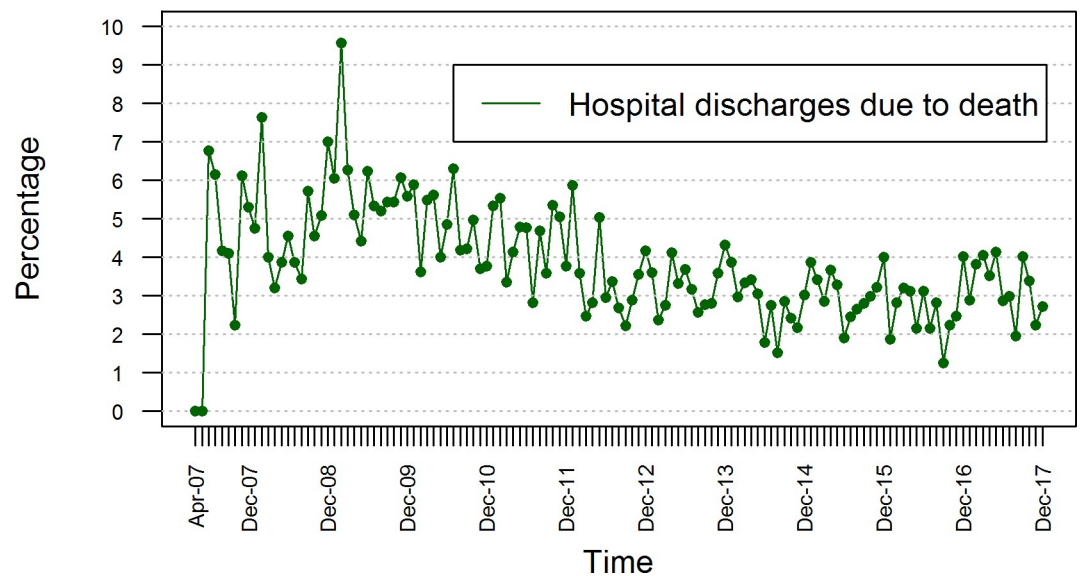
Year	Number of hospital discharges	Number of hospital discharges due to death	% Hospital discharges due to death
2007 (Apr-Dec)	623	29	4.65%
2008	3,045	146	4.79%
2009	5,063	292	5.77%
2010	5,443	257	4.72%
2011	5,464	244	4.47%
2012	5,884	201	3.42%
2013	6,361	206	3.24%
2014	7,246	198	2.73%
2015	7,867	240	3.05%
2016	8,555	225	2.63%
2017	8,949	289	3.23%
2018	8,843	221	2.50%
2019	7,740	219	2.83%
Total	81,083	2767	3.41%

<https://doi.org/10.1371/journal.pone.0248277.t001>

hospital discharges due to death, 5.77%, was recorded. In contrast, 2018 had the lowest annual percentage of hospital discharges due to death, 2.50%.

Fig 1 shows the time series of the monthly percentage of hospital discharges due to death in the HRAEB from April 2007 to December 2017. The series consists of 129 observations that oscillate between 0 and 10%, reaching their maximum at $t = 23$ (February 2009): 9.57%.

Fig 2A shows the graph of the estimated SLRMLS to model the trend of the percentage of hospital discharges due to death with respect to time, with $R^2 = 28.24\%$, slope = -0.021 (standard error (SE) = 0.003), and y-intercept = 5.182 (SE = 0.220). This model describes a trend with an apparent decrease in the percentage of hospital discharges due to death over time, but when verifying the assumptions of the model, it was observed that the residuals did not meet the assumption of independence ($p < 0.001$). This was done through the implementation of

**Fig 1. Time series of the monthly percentages of HRAEB hospital discharges due to death.**

<https://doi.org/10.1371/journal.pone.0248277.g001>

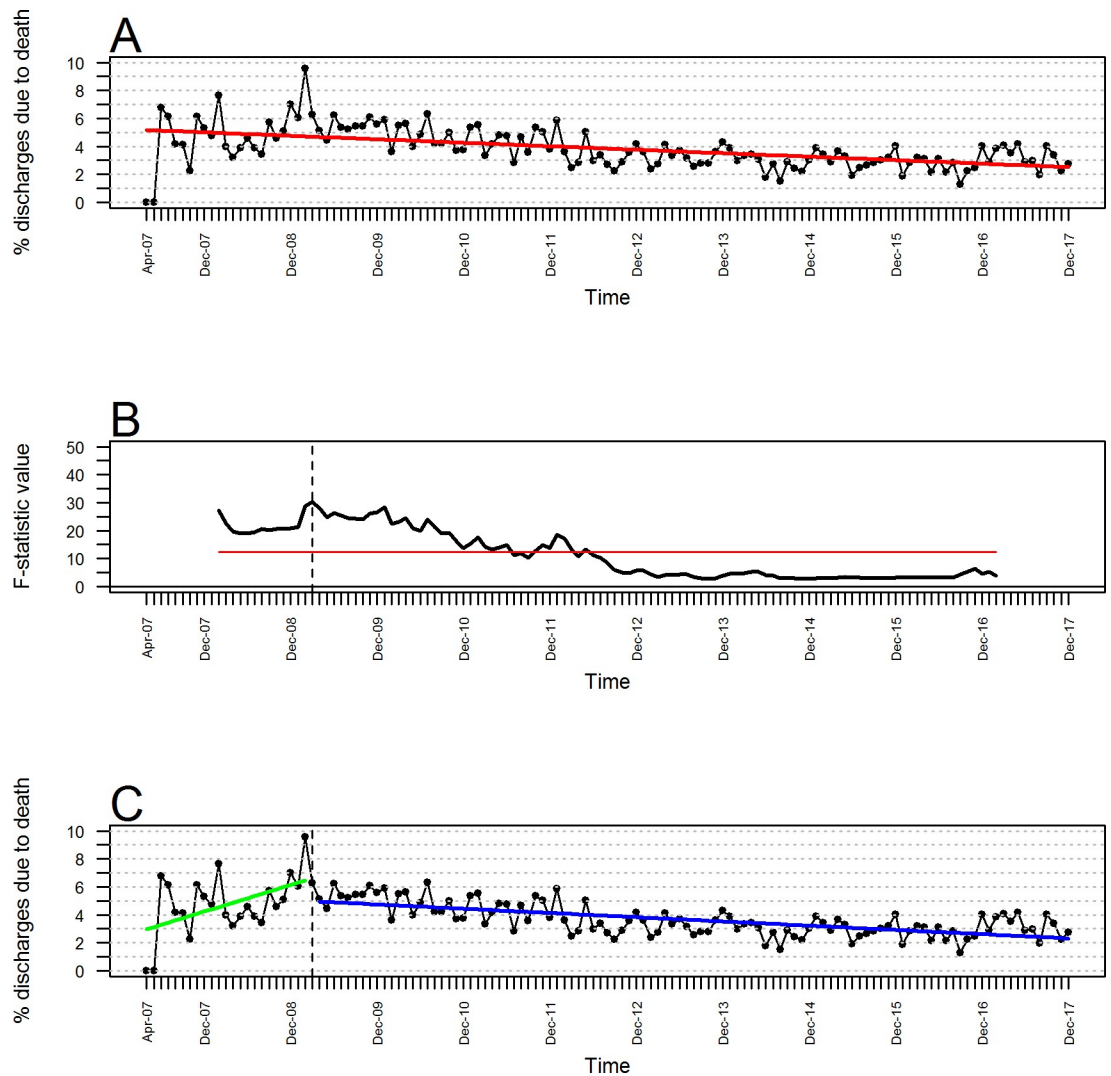


Fig 2. Trends and structural change in the monthly percentage of HRAEB hospital discharges due to death. (A) Trend modeling using a linear model. (B) F-statistics from the Chow test ($\alpha = 0.05$; red horizontal line). (C) Segmented trend modeling using two linear models.

<https://doi.org/10.1371/journal.pone.0248277.g002>

the Durbin-Watson test, which resulted in the detection of positive autocorrelation. Next, **Fig 2B** graphically illustrates the statistics associated with the Chow tests, which we performed to detect the possible instances where there was a structural change in the trend of the time series. There was a structural change at (March 2009) at the level. Finally, **Fig 2C** shows the graphs of the two estimated models (SLRMLS) that resulted from modeling the trend of the percentage of hospital discharges due to death with respect to time considering the structural change identified at $t = 24$. The first model describes the observations before $t = 24$; it has $R^2 = 24.00\%$, a slope of 0.159 (SE = 0.062), and a y-intercept of 2.804 (SE = 0.845). This model shows an increasing trend before $t = 24$. The second model describes the observations after $t = 24$; it has $R^2 = 44.46\%$, a slope of -0.025 (SE = 0.003) and a y-intercept of 5.565 (SE = 0.230). This model shows a decreasing trend after $t = 24$. Both models together show a significantly lower combined sum of squared estimate of errors (SSE; $SSE_1 + SSE_2$) than the SSE resulting from the linear-trend SLRMLS that described the decreasing trend in **Fig 2A** ($p < 0.001$).

Fig 3 shows the exploratory graph of the additive decomposition of the series of the monthly percentage of HRAEB hospital discharges due to death as a trend, from seasonal variations (seasonality), and from irregular fluctuations (random). An apparent increasing trend was observed from the beginning until mid-2009, and later, there was an apparent decreasing trend. There was also an apparent seasonal variation with period $s = 12$.

In trying to verify whether the series of the monthly percentage of hospital discharges due to death was stationary, the presence of a unit root could be identified (Dickey-Fuller = -3.4712 , lag order = 5, $p = 0.48$), from which we could presume that this was a nonstationary time series. When differentiating the series, we concluded that the series y_t was integrated of order $D = 1$, since it was necessary to differentiate the series once to obtain the series $(y_t - y_{t-1})$ that was stationary (Dickey-Fuller = -6.7907 , lag order = 5, $p < 0.01$).

Fig 4A shows the monthly series of hospital discharges due to death at the HRAEB (y_t). **Fig 4B** and **Fig 4C** show the autocorrelation (ACF) and partial autocorrelation (PACF) functions of the time series, respectively. The ACF function of the values of the time series was cut with extreme slowness, which reaffirmed that the series was nonstationary. Similarly, **Fig 5A** shows the differentiated series $(y_t - y_{t-1})$, and **Fig 5B** and **5C** show the ACF and PACF of the differentiated series, respectively. They were cut after the second phase, which reaffirmed that the differentiated series was stationary. In addition, the analysis of both figures empirically suggested that $p = 1$ or 2 and $q = 1$ or 2 may be relevant orders of the AR and MA components of an ARIMA model for the series studied. After confirming that the series of the monthly percentage of discharges due to death at the HRAEB could be modeled with a seasonal ARIMA model with integrated order $D = 1$, the AIC and BIC were calculated in models I (1), AR (0), AR (1), AR (2), MA (0), MA (1), and MA (2). The number of lags that minimized them occurred in the ARIMA(0,0,0)(1,1,2)₁₂ model, that is, when $P = 1, Q = 2$.

Based on the above, a seasonal ARIMA(0,0,0)(1,1,2)₁₂ with drift was adjusted to model the series of the monthly percentage of hospital discharges from HRAEB due to death without considering structural change. The coefficient associated with the AR component (1) was $\text{sar1} = 0.632$ (SE = 0.455), the coefficients associated with the MA component (2) were $\text{sma1} = -1.116$ (SE = 0.475) and $\text{sma2} = 0.231$ (SE = 0.280), and the drift coefficient was -0.017 (SE = 0.005). The AIC and BIC of this model were 450.52 and 419.33, respectively.

For the series of the monthly percentage of HRAEB hospital discharges due to death considering structural change, a seasonal ARIMA(0,0,0)(1,1,2)₁₂ with drift model was adjusted. The coefficient associated with the AR (1) component was $\text{sar1} = -0.991$ (SE = 0.116), the coefficients associated with the MA component (2) were $\text{sma1} = 0.289$ (SE = 0.275) and $\text{sma2} = -0.656$ (SE = 0.251), and the drift coefficient was -0.025 (SE = 0.004). The AIC and BIC of this model were 270.64 and 283.3, respectively.

Fig 6A shows the point predictions with prediction intervals of 90% of the percentage of hospital discharges due to death from the HRAEB for the 24 months included in the test dataset (January 2018 to December 2019) that resulted from adjusting the ARIMA(0,0,0)(1,1,2)₁₂ with drift, without considering a structural change. **Fig 6B** shows the predictions for the same period with prediction intervals of 90% that resulted with an ARIMA(0,0,0)(1,1,2)₁₂ model with drift but considering the structural change identified at $t = 24$. When comparing both predictions, the 90% prediction intervals that considered the structural change in the trend were narrower.

Table 2 shows the predictions for the monthly percentage of hospital discharges due to death from January 2018 to December 2019 considering the two models (no structural change and considering the structural change in the trend). The model that did not consider structural change estimated an average point prediction of the overall percentage of hospital discharges due to death of $2.62 \pm 0.39\%$ (range 1.73–3.32%) for the period from January 2018 to

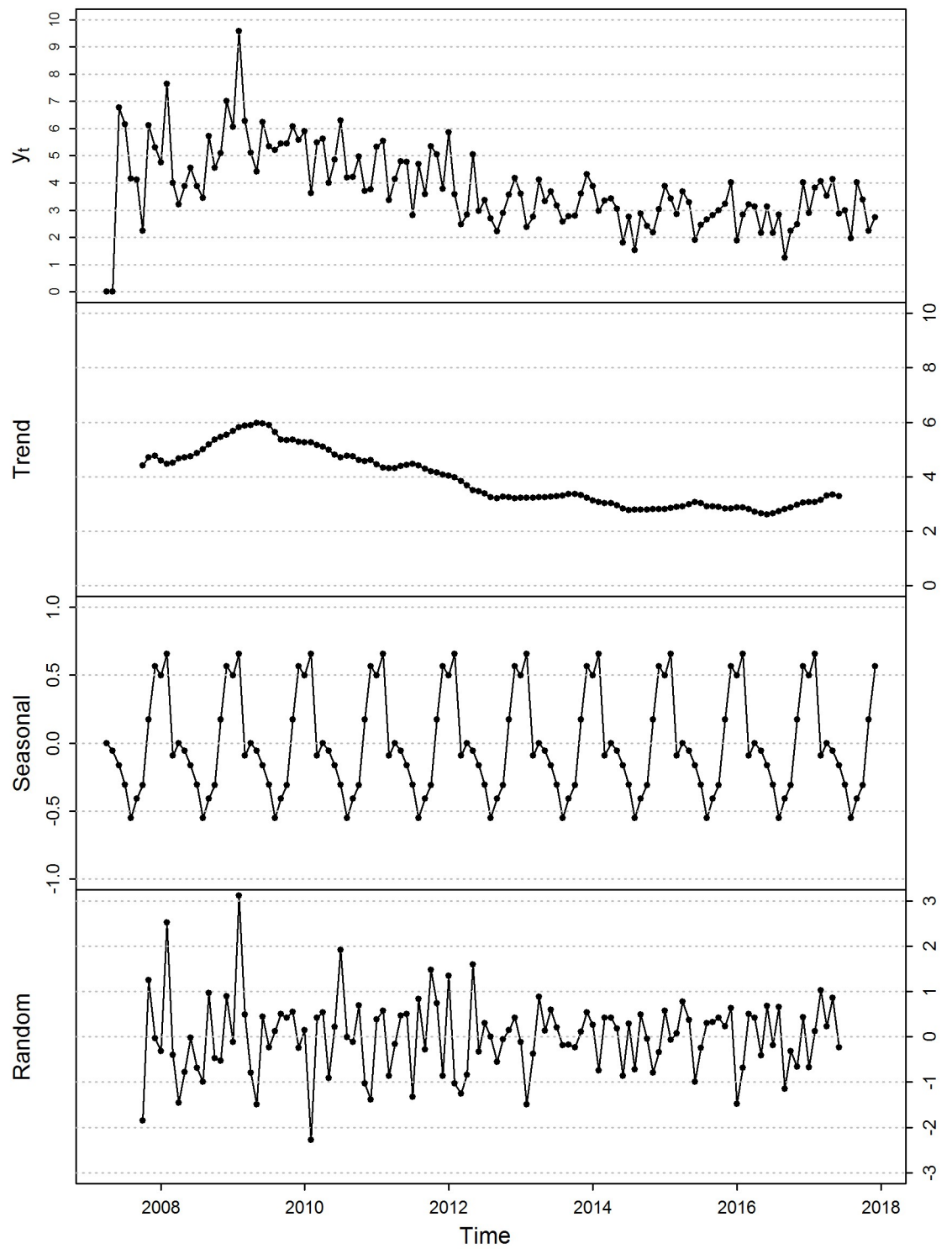


Fig 3. Additive decomposition of the monthly percentage of HRAEB hospital discharges due to death.

<https://doi.org/10.1371/journal.pone.0248277.g003>

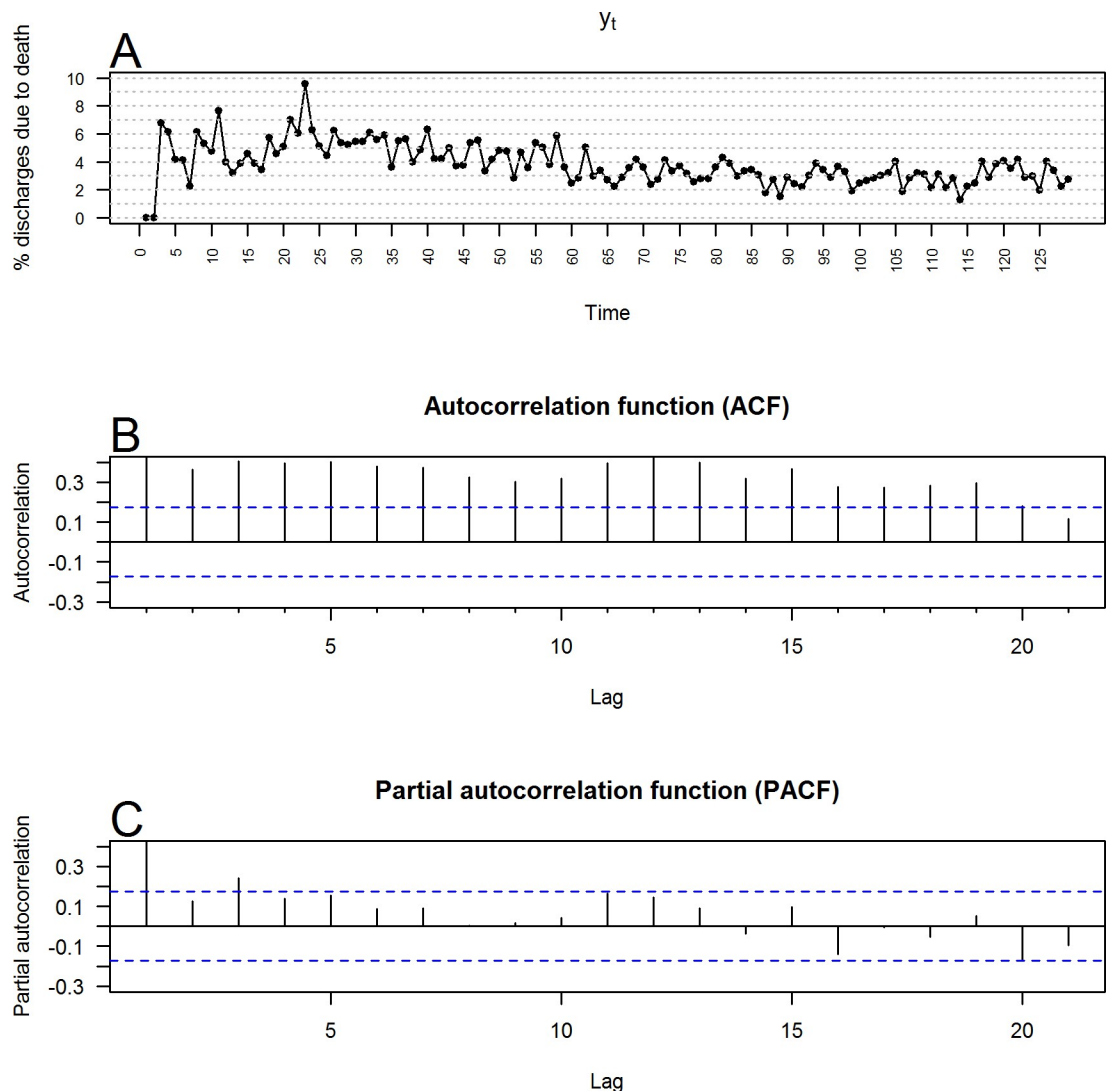


Fig 4. Monthly percentage of HRAEB hospital discharges due to death. (A) Original values. (B) Autocorrelation function ($\alpha = 0.05$; blue bands). (C) Partial autocorrelation function ($\alpha = 0.05$; blue bands).

<https://doi.org/10.1371/journal.pone.0248277.g004>

December 2019, a percentage higher than that estimated with the prediction average point generated by the model that incorporated the identified structural change $2.12 \pm 0.36\%$ (range 1.41–2.62%) for the same period. More accurate estimates (narrower prediction intervals) were obtained when using the model that incorporated the identified structural change, but compared with the actual values observed for the monthly percentage of hospital discharges due to death of $2.66 \pm 0.71\%$ (range 1.35–3.79%), very rough point estimates were observed in both cases. However, they were underestimated by the model that considered structural change.

In comparing the performance of the forecasts made for the monthly percentage of hospital discharges due to death using an $ARIMA(0,0,1)(1,1,2)_{12}$ with drift model that did not consider a structural change and an $ARIMA(0,0,0)(1,1,2)_{12}$ with drift model that did consider the structural change, the metrics listed in Table 3 were obtained, and all the calculated error indices favored the ARIMA model that did not consider the structural change, namely, $MAE = 0.63$, $RMSE = 0.81$, $MAPE = 25.89\%$, and $MASE = 0.65$.

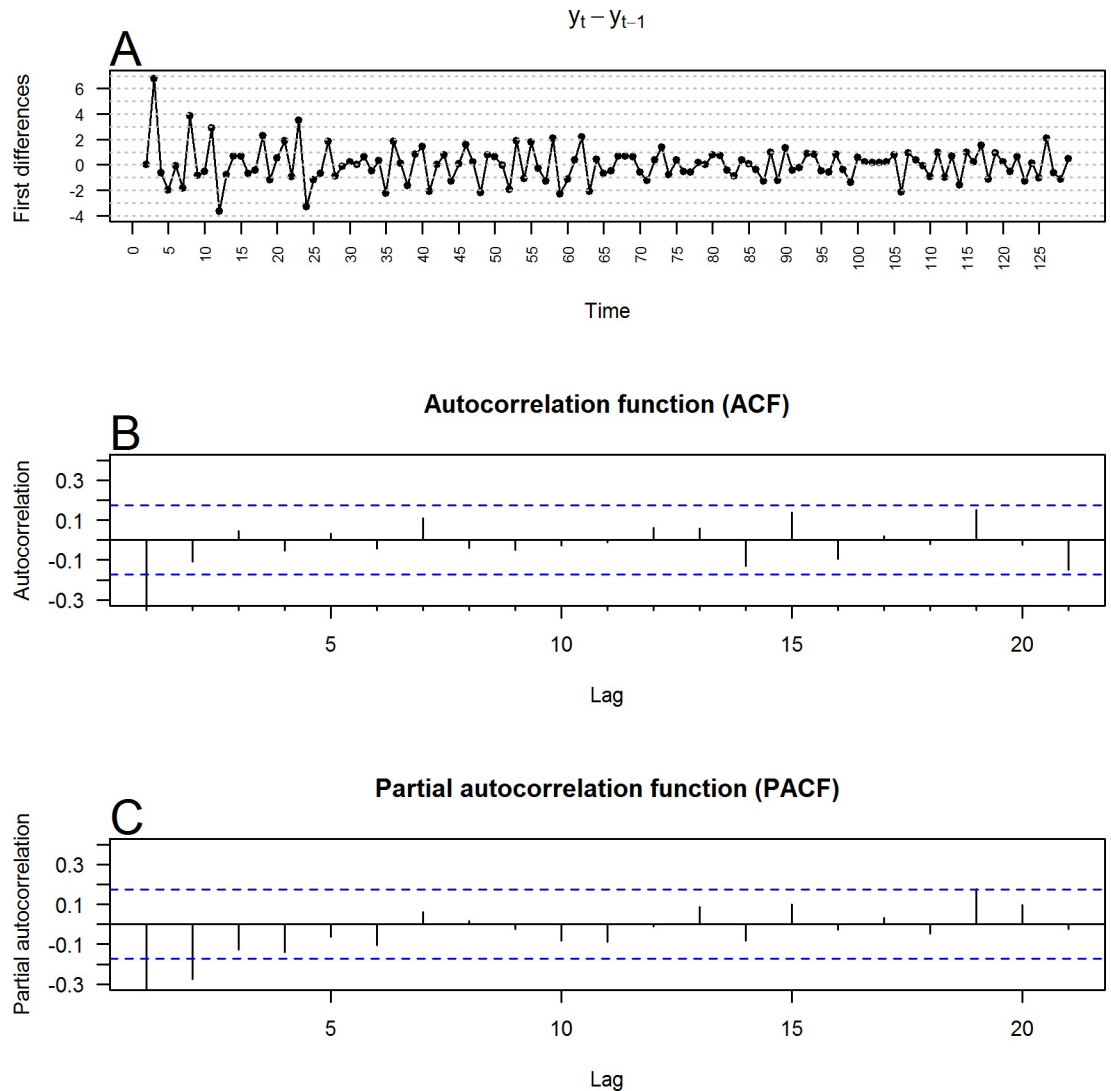


Fig 5. First differences in the monthly percentage of HRAEB hospital discharges due to death. (A) First differences. (B) Autocorrelation function ($\alpha = 0.05$; blue bands). (C) Partial autocorrelation function ($\alpha = 0.05$; blue bands).

<https://doi.org/10.1371/journal.pone.0248277.g005>

Discussion

In the study period (April 2007–December 2019), 3.41% (2,767/81,083) of hospital discharges from the HRAEB were due to death, which annually ranged from 2.50% to 5.77%, similar to the 4.10% identified by García. *et al.* [7] in a Spanish specialty hospital based on 24,194 episodes of hospitalization. These annual percentages are greater than the estimated 2% of hospital discharges per year due to death considering all Mexican SSA hospitals [3] and are also higher than the Mexican rate published in the journal *Salud Pública de México* in 1999 [2], which reported that 2.62% (35,891/1,469,161) of hospital discharges were due to death each year at the national level and 1.93% at the state level (Guanajuato, where the HRAEB is located) for 1999. There is little variability in the annual percentage of hospital discharges due to death at either the national or state level. The percentage of hospital discharges due to death at the HRAEB was higher than the national level and the state level (Guanajuato). This can be attributed to the fact that hospital discharges from the HRAEB correspond to patients with complex

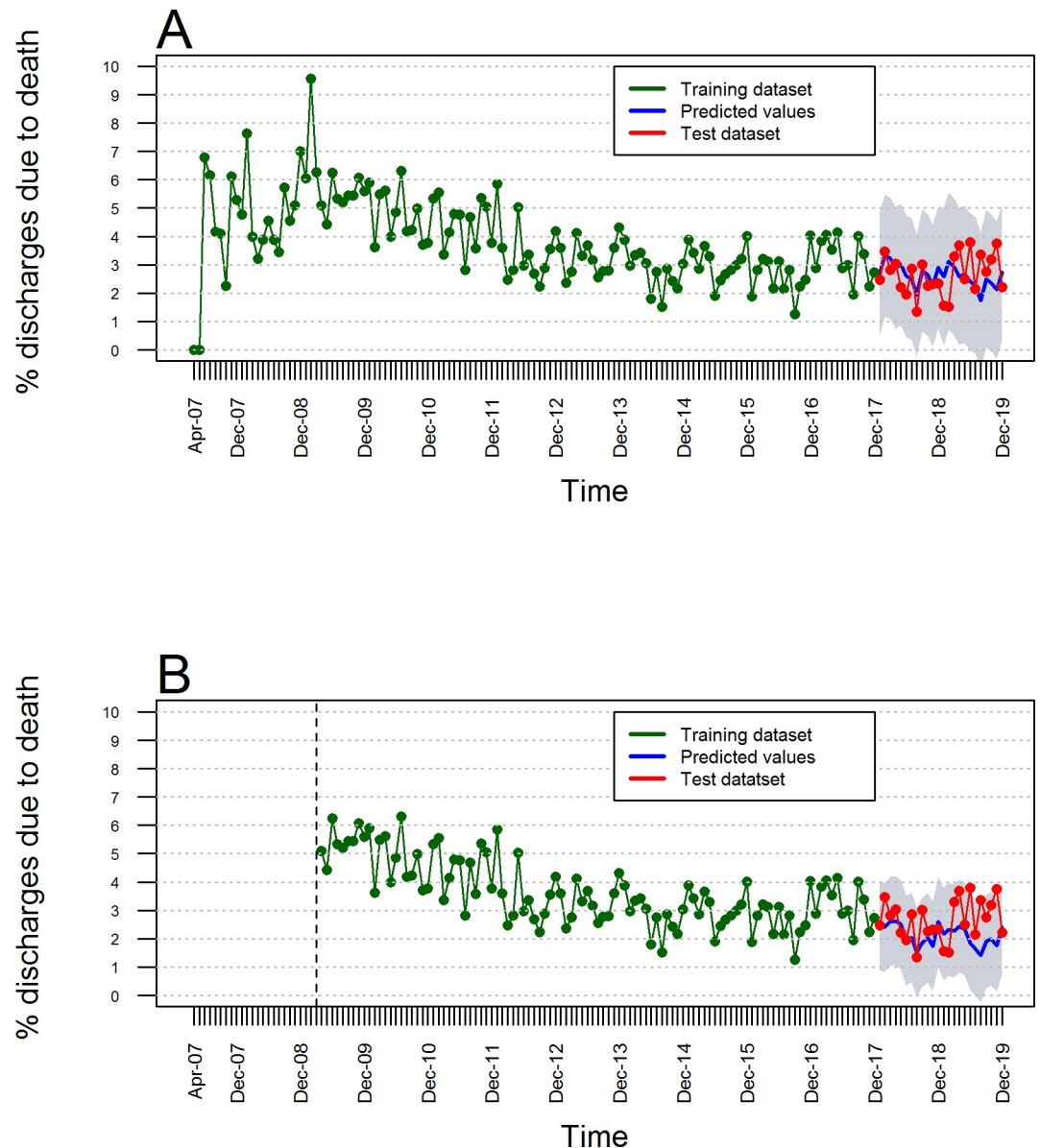


Fig 6. Forecasts from ARIMA(0,0,0)(1,1,2)₁₂ models with 90% prediction intervals. (A) Estimated predictions considering no structural change. (B) Estimated predictions considering the identified structural change at $t = 24$.

<https://doi.org/10.1371/journal.pone.0248277.g006>

pathologies that require tertiary care, as suggested by Jiménez [8] in their study on the quantification of the quality of hospital services.

Regarding the monthly time series of the percentage of hospital discharges due to death in the HRAEB, a sudden change (structural change) in the trend of the time series was identified; this statistically significant change ($p < 0.001$) was detected at $t = 24$ (March 2009), which suggests that the trend of the series can be modeled by two linear segments: one increasing from April 2007 to March 2009 and another decreasing from April 2009 to December 2017. The factors that could influence the increasing trend and subsequent decreasing trend in the percentage of HRAEB hospital discharges due to death are as follows: First, the gradual hiring of specialized nursing staff in the HRAEB. Each year, the number of nurses (certified and specialized) has increased, since, as described by Aiken *et al.* [22], the number of nursing personnel

Table 2. Predictions and observed values (test dataset) of the monthly percentage of HRAEB hospital discharges due to death.

t	Month	Model not considering structural change			Model considering the identified structural change (t = 24)			Observed values (test dataset)
		ARIMA(0,0,0)(1,1,2) ₁₂			ARIMA(0,0,0)(1,1,2) ₁₂			
		Prediction	90% prediction Interval		Prediction	90% prediction Interval		
130	Jan-18	2.70	0.55	4.85	2.48	0.90	4.05	2.48
131	Feb-18	3.32	1.17	5.47	2.40	0.82	3.98	3.48
132	Mar-18	3.23	1.08	5.38	2.61	1.04	4.19	2.82
133	Apr-18	2.89	0.74	5.04	2.62	1.04	4.20	3.03
134	May-18	3.01	0.86	5.16	2.51	0.93	4.09	2.20
135	Jun-18	2.60	0.45	4.74	1.93	0.35	3.51	1.95
136	Jul-18	2.49	0.34	4.64	2.05	0.47	3.63	2.87
137	Aug-18	1.92	-0.23	4.07	1.53	-0.05	3.11	1.35
138	Sep-18	2.81	0.66	4.96	1.87	0.29	3.45	3.01
139	Oct-18	2.64	0.49	4.79	2.03	0.45	3.61	2.25
140	Nov-18	2.25	0.10	4.40	1.74	0.16	3.32	2.33
141	Dec-18	2.90	0.75	5.05	2.61	1.03	4.19	2.35
142	Jan-19	2.58	0.16	5.00	2.16	0.52	3.80	1.56
143	Feb-19	3.13	0.71	5.55	2.33	0.69	3.97	1.53
144	Mar-19	2.92	0.50	5.34	2.28	0.64	3.92	3.30
145	Apr-19	2.59	0.17	5.01	2.43	0.80	4.07	3.68
146	May-19	2.67	0.25	5.09	2.35	0.71	3.99	2.50
147	Jun-19	2.40	-0.02	4.82	1.83	0.19	3.47	3.79
148	Jul-19	2.29	-0.13	4.71	1.65	0.01	3.29	2.15
149	Aug-19	1.73	-0.69	4.15	1.41	-0.23	3.05	3.37
150	Sep-19	2.50	0.08	4.92	1.91	0.27	3.55	2.76
151	Oct-19	2.37	-0.05	4.79	2.00	0.36	3.64	3.18
152	Nov-19	2.12	-0.30	4.54	1.77	0.13	3.41	3.75
153	Dec-19	2.74	0.32	5.16	2.41	0.77	4.05	2.21
	Average	2.62			2.12			2.66
	Std. Dev.	0.39			0.36			0.71
	Minimum	1.73			1.41			1.35
	Maximum	3.32			2.62			3.79

<https://doi.org/10.1371/journal.pone.0248277.t002>

as well as their level of professionalization affect hospital mortality. Leibson *et al.* [25] reaffirm the above facts and describe that a higher ratio of nurses per census bed decreases the mortality rate. Second, the constant increase in the volume of hospitalized patients in the first two years of HRAEB operation was associated with the habilitation and gradual certification of HRAEB

Table 3. Forecast-accuracy metrics for the monthly percentage of hospital discharges due to death at the HRAEB.

Type of index	Evaluated error rate	Monthly percentage of discharges due to death	
		ARIMA(0,0,0)(1,1,2) ₁₂ not considering structural change	ARIMA(0,0,0)(1,1,2) ₁₂ considering structural change (t = 24)
Scale-dependent errors	MAE	0.63	0.74
	RMSE	0.81	0.95
Percentage error	MAPE	25.89%	25.75%
Scale-free error	MASE	0.65	0.76

MAE: Mean absolute error, RMSE: Root mean square error, MAPE: Mean absolute percentage error, MASE: Mean absolute scaled error.

<https://doi.org/10.1371/journal.pone.0248277.t003>

services, consistent with the study by Taylor *et al.* [26], which described the association between the increase in the volume of patients and the increase in hospital mortality. Third, with the increase in the volume of surgical patients, approximately 80% of patients discharged from the HRAEB had undergone surgical procedures, which increases the likelihood of death from complex procedures, as detailed by Mcphee *et al.* [28], who identified the increase in the volume of surgeries as the main factor affecting mortality in one hospital.

The monthly predictions of the percentage of hospital discharges due to death from January 2018 to December 2019 were estimated by fitting seasonal ARIMA models, similar to the method of Fang *et al.* [12], but with and without consideration of the structural change identified by Chow's test [31]. In comparing the performance of the predictions of both models with the actual percentages of hospital discharges due to death in that period (test dataset), the four calculated error metrics (MAE, RMSE, MAPE, MASE) favored the predictions made with the model that did not consider structural change. In our models, the estimated MAPE errors of approximately 25–30% may seem high but are because the values of the percentage of discharges for death were small (0–10%), which indicates that MAPE is not a good criterion for choosing the best model in this case, as described by Hyndman and Koehler [40]. The MAE and RMSE error indices are frequently used to determine which model generates the best predictions, perhaps due to their ease of calculation and interpretation, as observed in the study by Liu *et al.* [43]. However, a study by Hyndman [39] suggests that MASE is the best error metric to evaluate predictions in time series, such as ours, detailing that MASE values less than 1 indicate more accurate predictions. In our study, the two models constructed to predict the percentage of discharges due to death generated MASE values less than 1, but the model that did not consider structural change had a lower MASE value, suggesting that it is a more accurate model.

In future work, our data will allow us to evaluate various types of models (cubic, ARIMA, joint point), as Fang *et al.* did [12], and it will be necessary to identify the services that have the greatest influence on hospital mortality. It is necessary to study the percentage of hospital discharges due to death, conditioned on medical services in the HRAEB, and perform an analysis similar to that of Andrews *et al.* [9] that conditions mortality by specific diagnoses and procedures. In addition, the trends in mortality related to the specializations offered at the HRAEB could be specified, similar to what was described by Gonzaga *et al.* [44] in patients with breast cancer, Segura *et al.* [11] in patients with tuberculosis, and Fang *et al.* [12] in patients with liver cancer, to name a few.

Finally, the study has several limitations. First, since this was a retrospective study, causality could not be inferred. Second, as a secondary database was used, there could be minor capture errors inherent to the source itself. Third, since it was a single-center study in a tertiary care hospital, the results cannot be generalized to hospitals of other levels of care, although the analysis method described for analyzing the time series of mortality could be implemented to identify structural changes in hospitals of different levels of care.

Conclusions

This study expands the knowledge on hospital mortality, showing that the overall percentage of hospital discharges due to death in a Mexican tertiary care hospital is 3.41%. Through the detection of a structural change (in March 2009) in the series of the monthly percentage of hospital discharges due to death, we identified that hospital mortality at the HRAEB can be described by a growing trend from its opening (April 2007) to March 2009 and by a decreasing trend from April 2009 to December 2019, in addition to seasonal behavior. Our study suggests that the seasonal ARIMA models constructed and evaluated with their respective prediction

intervals are an adequate tool for future monitoring of the monthly percentage of hospital discharges due to death, allowing us to detect observations that depart from the described trend and identify future structural changes.

Despite this was a retrospective single-center study based on a secondary database, all this information generated and its methodology can be used to improve decision-making and resource management to reduce the rate of deaths in both HRAEBs and other health institutions. Further research efforts could be also devoted to identifying the medical services that have the greatest influence on hospital mortality via various types of mathematical models.

Acknowledgments

The authors wish to thank all the patients and staff who participated in the study.

Author Contributions

Conceptualization: Edel Rafael Rodea-Montero, Rodolfo Guardado-Mendoza, Lina Sofía Palacio-Mejía.

Data curation: Edel Rafael Rodea-Montero, Brenda Jesús Rodríguez-Alcántar.

Formal analysis: Edel Rafael Rodea-Montero, Rodolfo Guardado-Mendoza, Brenda Jesús Rodríguez-Alcántar, Jesús Rubén Rodríguez-Núñez, Carlos Alberto Núñez-Colín, Lina Sofía Palacio-Mejía.

Investigation: Edel Rafael Rodea-Montero, Brenda Jesús Rodríguez-Alcántar, Jesús Rubén Rodríguez-Núñez, Carlos Alberto Núñez-Colín.

Methodology: Edel Rafael Rodea-Montero, Rodolfo Guardado-Mendoza, Lina Sofía Palacio-Mejía.

Project administration: Edel Rafael Rodea-Montero, Rodolfo Guardado-Mendoza, Lina Sofía Palacio-Mejía.

Software: Edel Rafael Rodea-Montero, Brenda Jesús Rodríguez-Alcántar, Jesús Rubén Rodríguez-Núñez, Carlos Alberto Núñez-Colín.

Supervision: Edel Rafael Rodea-Montero, Rodolfo Guardado-Mendoza, Lina Sofía Palacio-Mejía.

Validation: Edel Rafael Rodea-Montero, Rodolfo Guardado-Mendoza, Lina Sofía Palacio-Mejía.

Visualization: Edel Rafael Rodea-Montero, Brenda Jesús Rodríguez-Alcántar, Jesús Rubén Rodríguez-Núñez, Carlos Alberto Núñez-Colín.

Writing – original draft: Edel Rafael Rodea-Montero, Brenda Jesús Rodríguez-Alcántar, Jesús Rubén Rodríguez-Núñez, Carlos Alberto Núñez-Colín.

Writing – review & editing: Edel Rafael Rodea-Montero, Rodolfo Guardado-Mendoza, Lina Sofía Palacio-Mejía.

References

1. A SB. El análisis de la mortalidad hospitalaria como una medida de efectividad. *Med Clin.* 1988; 91: 139–41.
2. Estadística de egresos hospitalarios de la secretaría de salud, 1999. *Salud Publica Mex.* 2000; 42: 456–470. <https://doi.org/10.1590/S0036-36342001000500014> PMID: 11127077

3. DGIS. Bases de datos sobre egresos hospitalarios. 2020. Available: http://www.dgis.salud.gob.mx/contenidos/basesdedatos/da_egresoshosp_gobmx.html.
4. Secretaría de Salud de Aguascalientes. Egresos Hospitalarios 2010–2014. Available: <http://www.issea.gob.mx/egresos.asp>.
5. DGIS. Guía de intercambio de información sobre Reporte de información de egresos hospitalarios a la Secretaría De Salud. 2015. Available: <http://www.dgis.salud.gob.mx/contenidos/intercambio/egresos.html>.
6. Agency for Healthcare Research and Quality. AHRQ Quality Indicators—Guide to inpatient quality indicators: Quality of Care in Hospitals—Volume, Mortality, and Utilization. 2002; Revision 4.
7. García Ortega C, Almenara Barrios J GOJ. Tasas específicas de mortalidad en el Hospital de Algeciras durante el período 1995–1996. *Rev Esp Salud Pública*. 1997; 71: 305–15. PMID: [9445758](https://pubmed.ncbi.nlm.nih.gov/9445758/)
8. Jiménez Paneque RE. Indicadores de calidad y eficiencia de los servicios hospitalarios: una mirada actual. *Rev Cuba salud pública*. 2004; 30: 0–0.
9. Andrews RM, Russo C a, Pancholi M. Trends in Hospital Risk-Adjusted Mortality for Select Diagnoses and Procedures, 1994–2004: Statistical Brief #38. *Healthc Cost Util Proj Stat Briefs*. 2006; 1–7. Available: <http://www.ncbi.nlm.nih.gov/pubmed/21850773>. PMID: [21850773](https://pubmed.ncbi.nlm.nih.gov/21850773/)
10. Zeger SL, Irizarry R, Peng RD. On time series analysis of public health and biomedical data. *Annu Rev Public Health*. 2006; 27: 57–79. <https://doi.org/10.1146/annurev.publhealth.26.021304.144517> PMID: [16533109](https://pubmed.ncbi.nlm.nih.gov/16533109/)
11. Segura AM, Rey JJ, Arbelaéz MP. Tendencia de la mortalidad y los egresos hospitalarios por tuberculosis, antes y durante la implementación de la reforma del sector salud, Colombia, 1985–1999. 2004; 24: 115–123.
12. Fang J, Wu K, Zeng Y, Tang W, Du P, Xu X, et al. Liver Cancer Mortality Characteristics and Trends in China from 1991 to 2012. 2015; 16: 1959–1964.
13. Helfenstein U. *Statistical Methods in Medical Research* Box-Jenkins. 1996. <https://doi.org/10.1177/096228029600500102> PMID: [8743076](https://pubmed.ncbi.nlm.nih.gov/8743076/)
14. Nelson BK. *Statistical Methodology: V. Time Series Analysis Using Autoregressive Integrated Moving Average*. 1998; 5: 739–744. <https://doi.org/10.1111/j.1553-2712.1998.tb02493.x> PMID: [9678399](https://pubmed.ncbi.nlm.nih.gov/9678399/)
15. Otu A, George A. O, Jude O, Hope Ifeyinwa M, Andrew I. I. Application of Sarima Models in Modelling and Forecasting Nigeria's Inflation Rates. *Am J Appl Math Stat*. 2014; 2: 16–28. <https://doi.org/10.12691/ajams-2-1-4>
16. Roberts FS. *Discrete mathematical models, with applications to social, biological, and environmental problems*. 1976.
17. Hanke JE, Wichern D. *Business Forecasting: Pearson New International Edition*. Pearson Higher Ed; 2013.
18. Dong B, Ma X, Chen F. Investigating the Differences of Single-Vehicle and Multivehicle Accident Probability Using Mixed Logit Model. 2018; 2018.
19. Chen F, Chen S, Ma X. Crash Frequency Modeling Using Real-Time Environmental and Traffic Data and Unbalanced Panel Data Models. 2016; 1–16. <https://doi.org/10.3390/ijerph13060609> PMID: [27322306](https://pubmed.ncbi.nlm.nih.gov/27322306/)
20. Thirty day in-hospital mortality and the ' weekend effect.' 2010; 345: 5994. <https://doi.org/10.1111/j.1445-5994.2009.02067.x> PMID: [19811554](https://pubmed.ncbi.nlm.nih.gov/19811554/)
21. Mohammed MA, Sidhu KS, Rudge G, Stevens AJ. Weekend admission to hospital has a higher risk of death in the elective setting than in the emergency setting: a retrospective database study of national health service hospitals in England. *BMC Health Serv Res*. 2012; 12: 87. <https://doi.org/10.1186/1472-6963-12-87> PMID: [22471933](https://pubmed.ncbi.nlm.nih.gov/22471933/)
22. Aiken LH, Clarke SP, Sloane DM, Sochalski J. Hospital Nurse Staffing and Patient Mortality, Nurse Burnout, and Job Dissatisfaction. 2015; 288: 1987–1993.
23. Manuscript A, Patient S. NIH Public Access. 2011; 290: 1617–1623. <https://doi.org/10.1001/jama.290.12.1617.Educational>
24. Aiken L, Free E. Nurse staffing and education in Europe: if not now, when? 2014; 6736: 1789–1790. [https://doi.org/10.1016/S0140-6736\(14\)60188-4](https://doi.org/10.1016/S0140-6736(14)60188-4)
25. Leibson CL, Ph D, Stevens SR, Harris M, Ph D. Nurse Staffing and Inpatient Hospital Mortality. 2011; 1037–1045.
26. Taylor HD, Dennis DA, Crane HS. Relationship Between Mortality Rates and Hospital Patient Volume for Medicare Patients Undergoing Major Orthopaedic Surgery of the Hip, Knee, Spine, and Femur. 1997; 12.

27. Uk T, Staffing N, Group S. Patient volume, staffing, and workload in relation to risk-adjusted outcomes in a random stratified sample of UK neonatal intensive care units: a prospective evaluation. 2002; 359: 99–107.
28. Mcphee JT, Iii PR, Eslami MH, Arous EJ, Messina LM, Schanzer A. Surgeon case volume, not institution case volume, is the primary determinant of in-hospital mortality after elective open abdominal aortic aneurysm repair. *YMVA*. 2007; 53: 591–599.e2. <https://doi.org/10.1016/j.jvs.2010.09.063> PMID: [21144692](https://pubmed.ncbi.nlm.nih.gov/21144692/)
29. Dimick JB. Variation in Hospital Mortality Associated with Inpatient Surgery. 2009.
30. R Development Core Team. R: A language and environment, for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2006.
31. Chow GC. Tests of Equality Between Sets of Coefficients in Two Linear Regressions. *Econometrica*. 1960; 28: 591–605. <https://doi.org/10.2307/1910133>
32. Zeileis A, Leisch F, Hornik K, Kleiber C. strucchange: An R Package for Testing for Structural Change in Linear Regression Models. *J Stat Softw*. 2002; 1–38.
33. Box GEP, Jenkins GM, Reinsel GC, Ljung GM. Time series analysis: forecasting and control. John Wiley & Sons; 2015.
34. Hyndman Rob J., Khandakar Yeasmin. Automatic Time Series Forecasting: The forecast Package for R. *J Stat Softw*. 2008; 27: 22. Available: <http://www.jstatsoft.org/v27/i03/paper>.
35. Hyndman RJ, Athanasopoulos G, Bergmeir C, Caceres G, Chhay L, O'Hara-Wild M, et al. forecast: Forecasting functions for time series and linear models. 2018.
36. Makridakis S, Wheelwright SC, Hyndman RJ. Forecasting methods and applications. John Wiley & Sons; 2008.
37. Kriechbaumer T, Angus A, Parsons D, Rivas Casado M. An improved wavelet-ARIMA approach for forecasting metal prices. *Resour Policy*. 2014; 39: 32–41. <https://doi.org/10.1016/j.resourpol.2013.10.005>
38. Tylkowski J, Hojan M. Time decomposition and short-term forecasting of hydrometeorological conditions in the south Baltic Coastal zone of Poland. *Geosci*. 2019; 9. <https://doi.org/10.3390/geosciences9020068>
39. Hyndman R. Another Look at Forecast Accuracy Metrics for Intermittent Demand. *Foresight Int J Appl Forecast*. 2006; 4: 43–46.
40. Hyndman RJ, Koehler AB. Another look at measures of forecast accuracy. 2006; 22: 679–688. <https://doi.org/10.1016/j.ijforecast.2006.03.001>
41. Huang H, Song B, Xu P, Zeng Q, Lee J, Abdel-aty M. Macro and micro models for zonal crash prediction with application in hot zones identification. *JTRG*. 2016; 54: 248–256. <https://doi.org/10.1016/j.jtrangeo.2016.06.012>
42. Wen H, Zhang X, Zeng Q, Sze NN. Bayesian spatial-temporal model for the main and interaction effects of roadway and weather characteristics on freeway crash incidence. *Accid Anal Prev*. 2019; 132: 105249. <https://doi.org/10.1016/j.aap.2019.07.025> PMID: [31415995](https://pubmed.ncbi.nlm.nih.gov/31415995/)
43. Liu Q, Li Z, Ji Y, Martinez L, Zia UH, Javaid A, et al. Forecasting the seasonality and trend of pulmonary tuberculosis in Jiangsu Province of China using advanced statistical time-series analyses. *Infect Drug Resist*. 2019; 12: 2311–2322. <https://doi.org/10.2147/IDR.S207809> PMID: [31440067](https://pubmed.ncbi.nlm.nih.gov/31440067/)
44. Gonzaga CMR, Freitas-Junior R, Curado M-P, Sousa A-LL, Souza-Neto J-A, Souza MR. Temporal trends in female breast cancer mortality in Brazil and correlations with social inequalities: ecological time-series study. *BMC Public Health*. 2015; 15: 1–9. <https://doi.org/10.1186/1471-2458-15-1> PMID: [25563658](https://pubmed.ncbi.nlm.nih.gov/25563658/)