

# Tendencias, cambios estructurales y evaluación de modelos de series de tiempo para el pronóstico del alta hospitalaria por muerte en un hospital de tercer nivel en México

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

Publicado: 8 de marzo de 2021 • <https://doi.org/10.1371/journal.pone.0248277>

## Resumen

### Fondo

Los datos sobre altas hospitalarias pueden utilizarse como un instrumento valioso para la planificación y gestión hospitalaria. La cuantificación de las muertes puede considerarse una medida de la eficacia de la intervención hospitalaria, y un alto porcentaje de altas hospitalarias por muerte pueden estar asociadas a deficiencias en la calidad de la atención hospitalaria.

### Objetivo

Determinar el porcentaje global de altas hospitalarias por muerte en un hospital de tercer nivel en México desde su apertura, describir las características de la serie temporal generada a partir del porcentaje mensual de altas hospitalarias por muerte y realizar y evaluar predicciones.

### Métodos

Este fue un estudio retrospectivo que involucró los registros médicos de 81,083 pacientes que fueron dados de alta de un hospital de tercer nivel de atención desde abril de 2007 hasta diciembre de 2019 (primeros 153 meses de operación). Para el análisis y construcción de los modelos (dataset de entrenamiento) se utilizaron los registros de los primeros 129 meses (abril de 2007 a diciembre de 2017). Además, se utilizaron los registros de los últimos 24 meses (enero 2018 a diciembre 2019) para evaluar las predicciones realizadas (test dataset). Se identificó el cambio estructural (test de Chow), se ajustaron los modelos ARIMA, se estimaron las predicciones con y sin considerar el cambio estructural y se evaluaron las predicciones mediante índices de error (MAE, RMSE, MAPE y MASE).

### Resultados

El porcentaje total de altas por muerte fue del 3,41%. Se observó un cambio estructural en la serie temporal (marzo de 2009,  $p > 0,001$ ), y ARIMA(0,0,0)(1,1,2)<sub>12</sub> con modelos de deriva se ajustaron con y sin consideración del cambio estructural. Las métricas de error favorecieron al modelo que no consideró el cambio estructural (MAE = 0,63, RMSE = 0,81, MAPE = 25,89% y MASE = 0,65).

### Conclusión

Nuestro estudio sugiere que los modelos ARIMA son una herramienta adecuada para el seguimiento futuro del porcentaje mensual de altas hospitalarias por muerte, permitiéndonos detectar observaciones que se aparten de la tendencia descrita e identificar futuros cambios estructurales.

**Cita:** Rodea-Montero ER, Guardado-Mendoza R, Rodríguez-Alcántar BJ, Rodríguez-Núñez JR, Núñez-Colín CA, Palacio-Mejía LS (2021) Tendencias, cambios estructurales y evaluación de modelos de series de tiempo para pronosticar el alta hospitalaria anticipada hasta la muerte en un hospital de tercer nivel de México. PLoS ONE 16(3): e0248277. <https://doi.org/10.1371/journal.pone.0248277>

**Editor:** Qiang Zeng, Universidad Tecnológica del Sur de China, CHINA

**Recibido:** 15 de diciembre de 2020; **Aceptado:** 23 de febrero de 2021; **Publicado:** 8 de marzo de 2021

**Copyright:** © 2021 Rodea-Montero et al. Este es un artículo de acceso abierto distribuido bajo los términos de la licencia de atribución de [Creative Commons](#), que permite el uso, la distribución y la reproducción sin restricciones en cualquier medio, siempre que se acredeite al autor original y la fuente.

**Disponibilidad de datos:** Todos los datos que sustentan los hallazgos están disponibles en el sitio web de la Dirección General de Información en Salud de México en: [http://www.dgis.salud.gob.mx/contenidos/basesdedatos/da\\_egresoshosp\\_gobmx.html](http://www.dgis.salud.gob.mx/contenidos/basesdedatos/da_egresoshosp_gobmx.html).

**Financiamiento:** El/los autor(es) no recibieron financiamiento específico para este trabajo.

**Conflictos de intereses:** Los autores han declarado que no existen conflictos de intereses.

## Introducción

Los datos sobre altas hospitalarias por muerte pueden utilizarse como un instrumento valioso para la planificación y gestión hospitalaria [ 1 ]. En México, la Dirección General de Información en Salud (DGIS) es el órgano operativo de la Secretaría de Salud (SSA) que se encarga de generar estadísticas sobre salud y cuenta con varios subsistemas de información, entre ellos el Sistema Automatizado de Egresos Hospitalarios (SAEH) [ 2 ]. El número de altas hospitalarias nacionales registradas por los hospitales de la SSA mexicana como fuentes primarias en el SAEH se estima en 3 millones de casos por año, y de estos, aproximadamente el 2% son muertes [ 3 ]. Existen informes descriptivos sobre las altas hospitalarias nacionales y regionales [ 2 , 4 ], pero existen pocos estudios inferenciales sobre tendencias, identificación de variables asociadas o predicciones sobre dichas altas hospitalarias.

El Hospital Regional de Alta Especialidad del Bajío (HRAEB) brinda servicios clínicos, diagnósticos y de tercer tratamiento desde abril de 2007. Cuenta con 184 camas censales, no cuenta con servicio de urgencias y da de alta a pacientes con enfermedades complejas. Sus documentos fuente para el registro de las altas hospitalarias son el expediente clínico y el certificado de defunción. Los datos correspondientes a estos documentos son capturados en la plataforma informática del SAEH por personal del área de Estadísticas Médicas del HRAEB de acuerdo a los lineamientos de la DGIS [ 5 ].

Desde 1986, la Health Care Financing Administration ha incorporado la tasa de mortalidad hospitalaria como un comparador cualitativo de los hospitales estadounidenses, lo que ha fomentado el uso de indicadores de resultados de atención hospitalaria en todo el mundo [ 6 ]. La mortalidad hospitalaria es uno de los indicadores de calidad asistencial más utilizados [ 7 ], y la cuantificación de las muertes puede considerarse una medida de la eficacia de la intervención hospitalaria, aunque no hay que olvidar que en este indicador influyen otros factores, como la patología a tratar, la estructura de la población y la accesibilidad a la unidad hospitalaria [ 8 ]. Un alto porcentaje de altas hospitalarias por muerte pueden estar asociadas a deficiencias en la calidad de la atención hospitalaria [ 9 ].

Dado que en este caso el porcentaje de altas por muerte es una variación porcentual en el tiempo, se puede definir como una serie temporal. Para modelar esta serie de mortalidad, normalmente se construyen modelos lineales [ 10 ]. Algunos estudios incluso han analizado las tendencias de altas hospitalarias por una causa específica [ 11 ], y otros autores han estimado predicciones [ 12 ]. Por lo general, los modelos de promedio móvil integrado autorregresivo (ARIMA) se utilizan para analizar y pronosticar datos de series temporales y también son capaces de modelar una amplia gama de datos estacionales [ 13 – 15 ]. El modelo estadístico Box-Jenkins (ARIMA) es una técnica de pronóstico que ignora por completo las variables independientes al hacer el pronóstico. Toma en cuenta los datos históricos y los descompone en un proceso autorregresivo (AR), donde hay una memoria de eventos pasados; un proceso Integrado (I), que da cuenta de la estabilización o estacionariedad de los datos, facilitando la previsión; y un promedio móvil (MA) de los errores de pronóstico, de modo que cuanto más largos sean los datos históricos, más precisos serán los pronósticos, a medida que aprende con el tiempo [ 16 , 17 ].

Para la identificación del modelo ARIMA, dos estadísticos de bondad de ajuste que se utilizan más comúnmente para la selección del modelo son: Criterio de información de Akaike (AIC) y Criterio de información bayesiano de Schwarz (BIC). El AIC y BIC se determinan sobre la base de una función de probabilidad [ 15 ]. Cuando se comparan dos modelos, el que tiene menor AIC o BIC es generalmente mejor [ 18 ]. Los valores bajos de AIC o BIC sugieren que un modelo se adapta muy bien a los requisitos de bondad de ajuste y parsimonia [ 19 ].

Además, en cuanto a las muertes hospitalarias, existe evidencia científica de que determinados eventos, entre ellos la habilitación de ingresos hospitalarios los fines de semana [ 20 , 21 ], varían el número y el nivel educativo del personal de enfermería, así como la proporción de enfermeras por cama censal [ 22 – 25 ], aumentando el número de pacientes hospitalizados [ 26 , 27 ] y aumentando el volumen de pacientes quirúrgicos [ 28 , 29 ], puede aumentar el riesgo de mortalidad hospitalaria y por tanto elevar el porcentaje de altas hospitalarias por muerte.

El objetivo de este estudio fue determinar el porcentaje global de altas hospitalarias por muerte en un hospital de tercer nivel en México desde su apertura en abril de 2007 hasta diciembre de 2019, para describir las características de la serie de tiempo generada a partir del porcentaje mensual de altas hospitalarias por muerte, a muerte de abril de 2007 a diciembre de 2017 (primeros 129 meses de operación) y hacer y evaluar predicciones del porcentaje mensual de altas hospitalarias por muerte de enero de 2018 a diciembre de 2019 (24 meses después).

## Materiales y métodos

### Pacientes

Este estudio retrospectivo de pacientes hispano-mexicanos incluyó todos los registros de pacientes ( $n = 81,083$ ) que fueron dados de alta de un hospital de tercer nivel (HRAEB) ubicado en la ciudad de León en el estado de Guanajuato, México, desde su apertura en abril de 2007 a diciembre de 2019 (153 meses de operación). El conjunto de datos es secundario y se obtuvo de los registros del gobierno nacional del subsistema SAEH operado por la DGIS [ 3 ]. Este es un subsistema que recopila información de altas hospitalarias nacionales de hospitales mexicanos como fuentes primarias. El conjunto de datos se dividió en dos partes. Para el análisis y construcción de los modelos (dataset de entrenamiento) se utilizaron los registros de los primeros 129 meses (abril de 2007 a diciembre de 2017). Además, se utilizaron los registros de los últimos 24 meses (enero 2018 a diciembre 2019) para evaluar las predicciones realizadas (test dataset).

### Consideraciones éticas

El protocolo de este estudio fue revisado y aprobado por los Comités Institucionales de Investigación y Ética del HRAEB y el Instituto Nacional de Salud Pública de México (números de aprobación: CI/HRAEB/2019/046 y PT 211, respectivamente).

#### análisis estadístico

Todos los datos fueron analizados con el software estadístico R [ 30 ]. Inicialmente se implementó un análisis descriptivo y se graficó la serie temporal asociada al porcentaje mensual de altas hospitalarias por muerte. Para la serie temporal, la tendencia se modeló utilizando un modelo de regresión lineal simple utilizando mínimos cuadrados (SLRMLS) sin segmentación [ 10 ]. Los supuestos de normalidad, homocedasticidad e independencia de los residuales se verificaron mediante las pruebas de Kolmogorov-Smirnov, White y Durbin-Watson, respectivamente. A continuación, se identificaron las instancias donde podrían ocurrir posibles cambios estructurales (cambios repentinos en la tendencia de la serie temporal) mediante la prueba de Chow [ 31 ] con el paquete estadístico R "strucchange", que está diseñado para probar cambios estructurales en modelos de regresión lineal [ 32 ]. A continuación, se realizó una descomposición aditiva de la serie en sus componentes de la siguiente manera: tendencia, variaciones estacionales (estacionalidad) y fluctuaciones irregulares (aleatorias) [ 10 ]. Se utilizaron modelos ARIMA. El nombre de este procedimiento proviene de sus tres componentes: autorregresión (AR), integración (I) y media móvil (MA). El modelo ARIMA(p, d, q) se ajusta a los datos observados, donde p, d y q representan el orden de los tres componentes respectivos. En cuanto a los componentes integrados del modelo ARIMA, si el proceso estocástico asociado a una serie de tiempo tiene raíz unitaria, podemos concluir que se trata de una serie de tiempo no estacionaria. Por lo tanto, se implementó la prueba de Dickey-Fuller para cada una de las series  $y_t$  para determinar si era estacionaria. Si no, se procedió a diferenciar cada serie (generando una nueva serie  $z_t = y_t - y_{t-1}$  con  $t = 2, 3, \dots, n$ ) tantas veces como sea necesario (d) para obtener una serie temporal estacionaria y así concluir que la serie original  $y_t$  fue integrada de orden (d). Además, se construyeron gráficas de autocorrelación (ACF) y autocorrelación parcial (PACF) para cada serie y para aquellas series que resultaron de la diferenciación [ 33 ] para confirmar empíricamente si la serie podría considerarse estacionaria e identificar el posible número de rezagos para cada una de ellas. el resto de componentes del modelo ARIMA.

Para poder predecir el porcentaje mensual de altas por muerte, considerando y sin considerar el cambio estructural identificado en la serie, para el pronóstico se ajustaron los modelos ARIMA estacionales. El proceso estacional ARIMA(p, d, q)(P, D, Q) s viene dado por  $\Phi(B^s)\phi(B)(1-B^s)^D(1-B)^d y_t = c + \Theta(B^s)\theta(B)\epsilon_t$  donde  $\Phi(z)$  y  $\Theta(z)$  son polinomios de orden P y Q respectivamente, ninguno de los cuales contiene raíces dentro del círculo unitario. Si  $c \neq 0$ , hay un polinomio implícito de orden  $d + D$  en la función de pronóstico [ 34 ]. La selección del modelo ARIMA estacional adecuado en cada caso se realizó mediante la función auto.arima del paquete R "forecast" [ 35 ]. Esta función devuelve el mejor modelo ARIMA según el valor AIC o BIC [ 36 ]. La función realiza una búsqueda en un rango de posibles modelos dentro de las restricciones de orden y elige el más adecuado [ 37 ]. La función auto.arima se basa principalmente en el algoritmo Hyndman-Khandakar, que combina pruebas de elementos unitarios, minimización de pruebas de Akaike corregidas (AICc) y el método de máxima verosimilitud para obtener el modelo más apropiado para los datos. Se tuvieron en cuenta los criterios de bondad de ajuste, basados en el criterio de información [ 38 ]:  $AIC = -2\ln(L) + 2(p+q+P+Q+k)$  y  $BIC = -2\ln(L) + (p+q+P+Q+k)\ln(n)$  donde  $k = 1$  si  $c \neq 0$  y  $0$  en caso contrario, y  $L$  es la probabilidad maximizada del modelo ajustado a los datos diferenciados  $(1-B^s)^D(1-B)^d y_t$ . Luego de seleccionar los modelos para cada serie, se verificó que los residuales no tuvieran estructura de dependencia y que siguieran un proceso de ruido blanco.

Finalmente, se realizaron predicciones (pronósticos puntuales e intervalos de predicción) de la serie del porcentaje de altas por muerte, considerando y sin considerar el cambio estructural identificado en la serie utilizando la función de pronóstico del paquete de R "forecast" [ 35 ]. Estas predicciones se evaluaron utilizando cuatro índices de error: dos índices de error dependientes de la escala: error absoluto medio (MAE) y error cuadrático medio (RMSE); un índice de error porcentual: error porcentual absoluto medio (MAPE); y un índice de error sin escala: error escalado absoluto medio (MASE) [ 39 – 42 ]. En todas las pruebas se utilizó el nivel de significancia de  $\alpha = 0,05$ .

## Resultados

Un total de 81.083 altas hospitalarias, de las cuales 2.767 fueron defunciones, se generaron en el HRAEB desde abril de 2007 hasta diciembre de 2019 y se incluyeron en el análisis final de este estudio. En la tabla 1 se describen las altas hospitalarias totales del HRAEB por año, observándose un porcentaje total de altas por muerte del 3,41%. En 2009 se registró el mayor porcentaje anual de altas hospitalarias por muerte, 5,77%. En cambio, 2018 tuvo el menor porcentaje anual de altas hospitalarias por muerte, 2,50%.

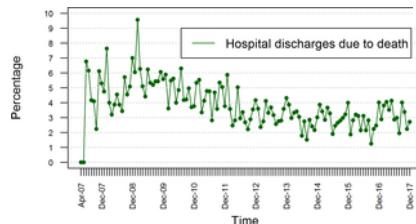
Year	Number of hospital discharges	Number of hospital discharges due to death	% Hospital discharges due to death
2007 (Apr-Dec)	623	29	4.65%
2008	3,045	146	4.79%
2009	5,063	292	5.77%
2010	5,441	257	4.72%
2011	5,861	244	4.17%
2012	5,881	201	3.42%
2013	6,261	206	3.24%
2014	7,246	198	2.73%
2015	7,867	240	3.05%
2016	8,551	231	2.63%
2017	8,949	289	3.23%
2018	8,843	221	2.50%
2019	7,740	219	2.83%
Total	81,083	2,767	3.41%

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

Tabla 1. Altas hospitalarias en el HRAEB agrupadas por año.

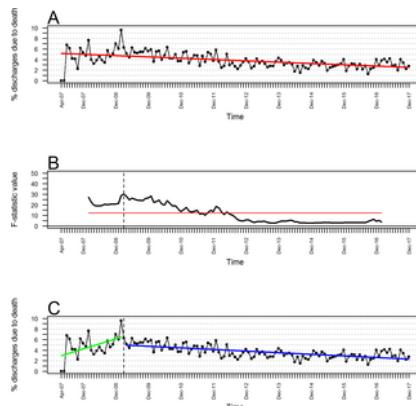
<https://doi.org/10.1371/journal.pone.0248277.t001>

La figura 1 muestra la serie temporal del porcentaje mensual de altas hospitalarias por muerte en el HRAEB desde abril de 2007 hasta diciembre de 2017. La serie consta de 129 observaciones que oscilan entre el 0 y el 10%, alcanzando su máximo en t = 23 (febrero de 2009). : 9,57%.



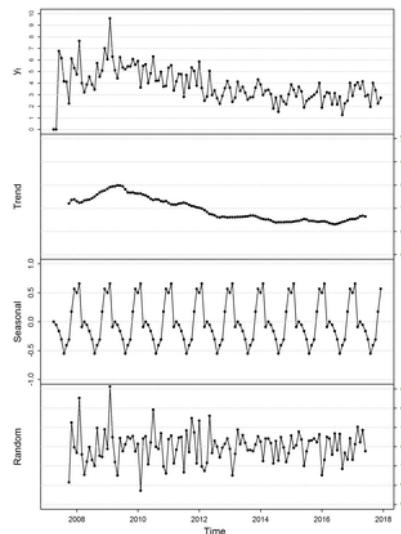
**Figura 1.** Serie temporal de los porcentajes mensuales de altas hospitalarias del HRAEB por fallecimiento.  
<https://doi.org/10.1371/journal.pone.0248277.g001>

**La figura 2A** muestra la gráfica del SLRMLS estimado para modelar la tendencia del porcentaje de altas hospitalarias por muerte con respecto al tiempo, con  $R^2 = 28,24\%$ , pendiente = -0,021 (error estándar (SE) = 0,003), y intersección y = 5,182 (SE = 0,220). Este modelo describe una tendencia con una aparente disminución del porcentaje de egresos hospitalarios por muerte a lo largo del tiempo, pero al verificar los supuestos del modelo se observó que los residuales no cumplían el supuesto de independencia ( $p < 0,001$ ). Esto se hizo mediante la implementación de la prueba de Durbin-Watson, que resultó en la detección de autocorrelación positiva. A continuación, **figura 2B** ilustra gráficamente las estadísticas asociadas con las pruebas de Chow, que realizamos para detectar las posibles instancias donde hubo un cambio estructural en la tendencia de la serie temporal. Hubo un cambio estructural en (marzo de 2009) a nivel. Finalmente, la **Fig. 2C** muestra las gráficas de los dos modelos estimados (SLRMLS) que resultaron de modelar la tendencia del porcentaje de altas hospitalarias por muerte con respecto al tiempo considerando el cambio estructural identificado en  $t = 24$ . El primer modelo describe las observaciones antes de  $t = 24$ ; tiene  $R^2 = 24,00\%$ , una pendiente de 0,159 (SE = 0,062) y una intersección en y de 2,804 (SE = 0,845). Este modelo muestra una tendencia creciente antes de  $t = 24$ . El segundo modelo describe las observaciones después de  $t = 24$ , tiene  $R^2 = 44,46\%$ , una pendiente de -0,025 (SE = 0,003) y un intercepto en y de 5,565 (SE = 0,230). Este modelo muestra una tendencia decreciente después de  $t = 24$ . Ambos modelos juntos muestran una suma combinada de estimaciones de errores al cuadrado significativamente menor (SSE;  $SSE_1 + SSE_2$ ) que el SSE resultante del SLRMLS de tendencia lineal que describió la tendencia decreciente en **Figura 2A** ( $p < 0,001$ ).



**Figura 2.** Evolución y cambio estructural del porcentaje mensual de altas hospitalarias del HRAEB por fallecimiento.  
(A) Modelado de tendencias utilizando un modelo lineal. (B) Estadísticas F de la prueba de Chow ( $\alpha = 0,05$ ; línea horizontal roja). (C) Modelado de tendencias segmentadas utilizando dos modelos lineales.  
<https://doi.org/10.1371/journal.pone.0248277.g002>

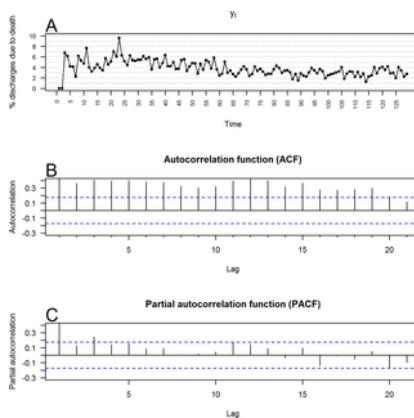
**La figura 3** muestra el gráfico exploratorio de la descomposición aditiva de la serie del porcentaje mensual de altas hospitalarias por muerte del HRAEB como tendencia, por variaciones estacionales (estacionalidad) y por fluctuaciones irregulares (aleatorias). Se observó una aparente tendencia creciente desde el inicio hasta mediados de 2009 y, posteriormente, una aparente tendencia decreciente. También hubo una aparente variación estacional con el período s = 12.



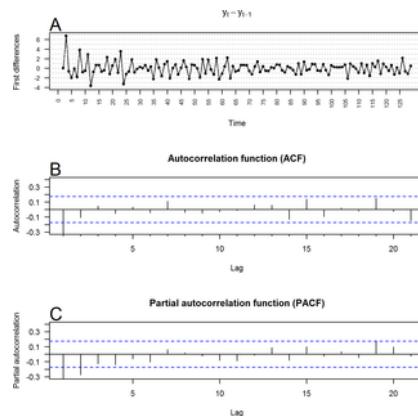
**Figura 3. Descomposición aditiva del porcentaje mensual de altas hospitalarias del HRAEB por fallecimiento.**  
<https://doi.org/10.1371/journal.pone.0248277.g003>

Al tratar de verificar si la serie del porcentaje mensual de altas hospitalarias por muerte era estacionaria, se pudo identificar la presencia de una raíz unitaria (Dickey-Fuller = -3.4712, orden de desfase = 5,  $p = 0.48$ ), a partir de la cual podría suponer que se trata de una serie de tiempo no estacionaria. Al diferenciar la serie, concluimos que la serie  $y_t$  estaba integrada de orden  $D = 1$ , ya que era necesario diferenciar la serie una vez para obtener la serie ( $y_t - y_{t-1}$ ) que era estacionaria (Dickey-Fuller = -6,7907, orden de retardo = 5,  $p < 0.01$ ).

**La figura 4A** muestra la serie mensual de altas hospitalarias por fallecimiento en el HRAEB ( $y_t$ ). **La figura 4B y la figura 4C** muestran las funciones de autocorrelación (ACF) y autocorrelación parcial (PACF) de la serie temporal, respectivamente. La función ACF de los valores de la serie temporal se cortó con extrema lentitud, lo que reafirmó que la serie no era estacionaria. De manera similar, **la Fig. 5A** muestra la serie diferenciada ( $y_t - y_{t-1}$ ), y **las Fig. 5B y 5C** muestran los ACF y PACF de las series diferenciadas, respectivamente. Se cortaron después de la segunda fase, que reafirmó que la serie diferenciada era estacionaria. Además, el análisis de ambas figuras sugirió empíricamente que  $p = 1$  o  $2$  y  $q = 1$  o  $2$  pueden ser órdenes relevantes de los componentes AR y MA de un modelo ARIMA para la serie estudiada. Después de comprobar que la serie del porcentaje mensual de altas por muerte en el HRAEB se podía modelar con un modelo ARIMA estacional con orden integrado  $D = 1$ , se calcularon los AIC y BIC en los modelos I (1), AR (0), AR (1), AR (2), MA (0), MA (1) y MA (2). El número de rezagos que los minimizó ocurrió en el modelo ARIMA(0,0,0)(1,1,2)<sub>12</sub>, es decir, cuando  $P = 1, Q = 2$ .



**Figura 4. Porcentaje mensual de altas hospitalarias del HRAEB por fallecimiento.**  
(A) Valores originales. (B) Función de autocorrelación ( $\alpha = 0.05$ ; bandas azules). (C) Función de autocorrelación parcial ( $\alpha = 0.05$ ; bandas azules).  
<https://doi.org/10.1371/journal.pone.0248277.g004>



**Figura 5.** Primeras diferencias en el porcentaje mensual de altas hospitalarias del HRAEB por fallecimiento.

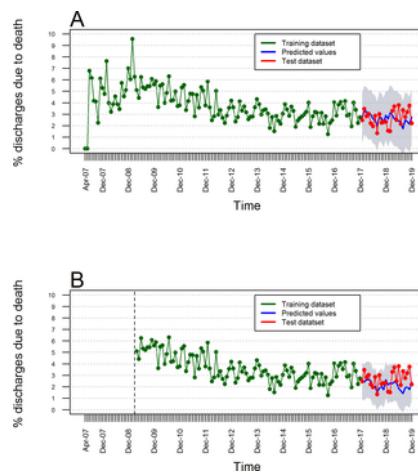
(A) Primeras diferencias. (B) Función de autocorrelación ( $\alpha = 0,05$ ; bandas azules). (C) Función de autocorrelación parcial ( $\alpha = 0,05$ ; bandas azules).

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

Con base en lo anterior, se ajustó un ARIMA(0,0,0)(1,1,2)<sub>12</sub> estacional con deriva para modelar la serie del porcentaje mensual de altas hospitalarias del HRAEB por muerte sin considerar cambio estructural. El coeficiente asociado al componente AR (1) fue sar1 = 0.632 (SE = 0.455), los coeficientes asociados al componente MA (2) fueron sma1 = -1.116 (SE = 0.475) y sma2 = 0.231 (SE = 0.280), y el coeficiente de deriva fue -0,017 (SE = 0.005). El AIC y BIC de este modelo fueron 450,52 y 419,33, respectivamente.

Para la serie del porcentaje mensual de altas hospitalarias HRAEB por muerte considerando el cambio estructural, se ajustó un ARIMA(0,0,0)(1,1,2)<sub>12</sub> estacional con modelo de deriva. El coeficiente asociado al componente AR (1) fue sar1 = -0.991 (SE = 0.116), los coeficientes asociados al componente MA (2) fueron sma1 = 0.289 (SE = 0.275) y sma2 = -0.656 (SE = 0.251), y el coeficiente de deriva fue -0,025 (SE = 0.004). El AIC y BIC de este modelo fueron 270,64 y 283,3, respectivamente.

**La figura 6A** muestra las predicciones puntuales con intervalos de predicción del 90 % del porcentaje de altas hospitalarias por muerte del HRAEB para los 24 meses incluidos en el conjunto de datos de prueba (enero de 2018 a diciembre de 2019) que resultaron de ajustar el ARIMA(0,0,0)(1,1,2)<sub>12</sub> con deriva, sin considerar un cambio estructural. **La Fig. 6B** muestra las predicciones para el mismo periodo con intervalos de predicción del 90% que resultaron con un modelo ARIMA(0,0,0)(1,1,2)<sub>12</sub> con deriva pero considerando el cambio estructural identificado en t = 24. Cuando comparando ambas predicciones, los intervalos de predicción del 90% que consideraron el cambio estructural en la tendencia fueron más estrechos.



**Fig 6.** Pronósticos de ARIMA(0,0,0)(1,1,2)<sub>12</sub> modelos con intervalos de predicción del 90%.

(A) Predicciones estimadas considerando ningún cambio estructural. (B) Predicciones estimadas considerando el cambio estructural identificado en t = 24.

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

**Tabla 2** muestra las predicciones para el porcentaje mensual de altas hospitalarias por muerte desde enero de 2018 hasta diciembre de 2019 considerando los dos modelos (sin cambio estructural y considerando el cambio estructural en la tendencia). El modelo que no consideró el cambio estructural estimó una predicción puntual promedio del porcentaje global de altas hospitalarias por muerte de  $2,62 \pm 0,39\%$  (rango 1,73-3,32%) para el período de enero de 2018 a diciembre de 2019, porcentaje superior al

estimado con el punto promedio de predicción generado por el modelo que incorporó el cambio estructural identificado  $2.12 \pm 0.36\%$  (rango 1.41–2.62%) para el mismo período. Se obtuvieron estimaciones más precisas (intervalos de predicción más estrechos) al usar el modelo que incorporó el cambio estructural identificado, pero en comparación con los valores reales observados para el porcentaje mensual de altas hospitalarias por muerte de  $2.66 \pm 0.71\%$  (rango 1.35–3.79%), se observaron estimaciones puntuales muy aproximadas en ambos casos. Sin embargo, fueron subestimados por el modelo que consideraba el cambio estructural.

t	Month	Model not considering structural change		Model considering the idea of structural change		Observed values (test dataset)
		ARIMA(0,0,0)(1,1,2) <sub>12</sub>	ARIMA(0,0,0)(1,1,2) <sub>12</sub>	ARIMA(0,0,0)(1,1,2) <sub>12</sub>	ARIMA(0,0,0)(1,1,2) <sub>12</sub>	
120	Jan-18	2.70	2.66	4.00	0.96	3.00
121	Feb-18	3.12	3.17	3.47	2.40	3.00
122	Mar-18	3.23	3.08	3.58	2.61	3.19
123	Apr-18	2.89	3.74	3.64	2.42	3.64
124	May-18	3.03	3.00	3.51	2.81	3.00
125	Jun-18	2.40	0.45	4.74	1.93	3.51
126	Jul-18	2.49	3.34	4.64	2.01	3.63
127	Aug-18	3.02	0.70	4.33	1.33	3.35
128	Sep-18	2.81	0.66	4.96	1.87	3.45
129	Oct-18	2.64	0.49	4.79	2.01	3.61
130	Nov-18	2.23	0.49	4.93	1.74	3.53
131	Dec-18	2.90	0.73	5.03	2.61	4.19
132	Jan-19	2.38	0.16	5.00	2.10	3.52
133	Feb-19	3.13	0.13	5.23	2.00	3.55
134	Mar-19	2.82	0.30	5.24	2.28	3.82
135	Apr-19	2.39	0.17	5.01	2.43	3.60
136	May-19	2.67	0.20	5.00	2.37	3.59
137	Jun-19	2.40	-0.62	4.82	1.83	3.47
138	Jul-19	2.29	-0.13	4.71	1.65	3.01
139	Aug-19	2.73	0.09	4.71	1.51	3.25
140	Sep-19	2.30	0.08	4.92	1.93	3.27
141	Oct-19	2.37	-0.05	4.79	2.00	3.64
142	Nov-19	2.12	0.08	4.71	1.77	3.11
143	Dec-19	2.74	0.32	5.16	2.61	3.79
Average		2.42			3.12	
MAE		0.63			0.66	
MSE		0.20			0.21	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	
MAE		0.63			0.66	
MAPE		25.89%			25.77%	
Scale-free error		MAE			0.65	
Scale-free error		MAPE			0.76	

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.

## References

1. A SB. El análisis de la mortalidad hospitalaria como una medida de efectividad. *Med Clin.* 1988;91: 139–41.  
[View Article](#) • [Google Scholar](#)
2. Estadística de egresos hospitalarios de la secretaría de salud, 1999. *Salud Publica Mex.* 2000;42: 456–470. pmid:11127077  
[View Article](#) • [PubMed/NCBI](#) • [Google Scholar](#)
3. DGIS. Bases de datos sobre egresos hospitalarios. 2020. Available:  
[http://www.dgis.salud.gob.mx/conteidos/basesdedatos/da\\_egresoshosp\\_gobmx.html](http://www.dgis.salud.gob.mx/conteidos/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/conteidos/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 periodo 1995–1996. *Rev Esp Salud Pública.* 1997;71: 305–15. pmid:9445758  
[View Article](#) • [PubMed/NCBI](#) • [Google Scholar](#)
8. Jiménez Paneque RE. Indicadores de calidad y eficiencia de los servicios hospitalarios: una mirada actual. *Rev Cuba salud pública.* 2004;30: 0–0.  
[View Article](#) • [Google Scholar](#)
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

[View Article](#) • [PubMed/NCBI](#) • [Google Scholar](#)

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. pmid:16533109  
[View Article](#) • [PubMed/NCBI](#) • [Google Scholar](#)
11. Segura AM, Rey JJ, Arbeláez 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.  
[View Article](#) • [Google Scholar](#)
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.  
[View Article](#) • [Google Scholar](#)
13. Helfenstein U. Statistical Methods in Medical Research Box-Jenkins. 1996. pmid:8743076  
[View Article](#) • [PubMed/NCBI](#) • [Google Scholar](#)
14. Nelson BK. Statistical Methodology: V. Time Series Analysis Using Autoregressive Integrated Moving Average. *1998;5:* 739–744. pmid:9678399  
[View Article](#) • [PubMed/NCBI](#) • [Google Scholar](#)
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.  
[View Article](#) • [Google Scholar](#)
16. Roberts FS. Discrete mathematical models, with applications to social, biological, and environmental problems. 1976.  
[View Article](#) • [Google Scholar](#)
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.*  
[View Article](#) • [Google Scholar](#)
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.* pmid:27322306  
[View Article](#) • [PubMed/NCBI](#) • [Google Scholar](#)
20. Thirty day in-hospital mortality and the 'weekend effect.' *2010;345:* 5994. pmid:19811554  
[View Article](#) • [PubMed/NCBI](#) • [Google Scholar](#)
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. pmid:22471933  
[View Article](#) • [PubMed/NCBI](#) • [Google Scholar](#)
22. Aiken LH, Clarke SP, Sloane DM, Sochalski J. Hospital Nurse Staffing and Patient Mortality, Nurse Burnout, and Job Dissatisfaction. *2015;288:* 1987–1993.  
[View Article](#) • [Google Scholar](#)
23. Manuscript A, Patient S. NIH Public Access. *2011;290:* 1617–1623.  
[View Article](#) • [Google Scholar](#)
24. Aiken L, Free E. Nurse staffing and education in Europe: if not now, when? *2014;6736:* 1789–1790.  
[View Article](#) • [Google Scholar](#)
25. Leibson CL, Ph D, Stevens SR, Harris M, Ph D. Nurse Staffing and Inpatient Hospital Mortality. *2011; 1037–1045.*  
[View Article](#) • [Google Scholar](#)
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.*  
[View Article](#) • [Google Scholar](#)
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.  
[View Article](#) • [Google Scholar](#)

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. pmid:21144692  
[View Article](#) • [PubMed/NCBI](#) • [Google Scholar](#)
29. Dimick JB. Variation in Hospital Mortality Associated with Inpatient Surgery. 2009.  
[View Article](#) • [Google Scholar](#)
30. R Development Core Team. R: A language and enviroment, 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.  
[View Article](#) • [Google Scholar](#)
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.  
[View Article](#) • [Google Scholar](#)
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.stats.ox.ac.uk/pub/MASS4>  
[View Article](#) • [Google Scholar](#)
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.  
[View Article](#) • [Google Scholar](#)
36. Makridakis S, Wheelwright SC, Hyndman RJ. Forecasting methods and applications. John wiley & sons; 2008.
37. Kriegbaumer T, Angus A, Parsons D, Rivas Casado M. An improved wavelet-ARIMA approach for forecasting metal prices. *Resour Policy*. 2014;39: 32–41.  
[View Article](#) • [Google Scholar](#)
38. Tykłowski J, Hojan M. Time decomposition and short-term forecasting of hydrometeorological conditions in the south Baltic Coastal zone of Poland. *Geosci*. 2019;9.  
[View Article](#) • [Google Scholar](#)
39. Hyndman R. Another Look at Forecast Accuracy Metrics for Intermittent Demand. *Foresight Int J Appl Forecast*. 2006;4: 43–46.  
[View Article](#) • [Google Scholar](#)
40. Hyndman RJ, Koehler AB. Another look at measures of forecast accuracy. 2006;22: 679–688.  
[View Article](#) • [Google Scholar](#)
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.  
[View Article](#) • [Google Scholar](#)
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. pmid:31415995  
[View Article](#) • [PubMed/NCBI](#) • [Google Scholar](#)
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. pmid:31440067  
[View Article](#) • [PubMed/NCBI](#) • [Google Scholar](#)
44. Gonzaga CMR, Freitas-Junior R, Curado MP, Sousa A-LL, Souza-Neto JA, Souza MR. Tendencias temporales en la mortalidad por cáncer de mama femenino en Brasil y correlaciones con las desigualdades sociales: estudio de series temporales ecológicas. *BMC Salud Pública*. 2015;15: 1–9. medio:25563658  
[Ver artículo](#) • [PubMed/NCBI](#) • [Google Académico](#)