



# Uso del enfoque bayesiano para la estimación y proyección de tasas

Nora Arnesi,<sup>1</sup> Leticia Hachuel<sup>1</sup> y Marcos Prunello<sup>1</sup>

## Forma de citar

Arnesi N, Hachuel L, Prunello M. Uso del enfoque bayesiano para la estimación y proyección de tasas. Rev Panam Salud Publica. 2015;38(4):286-91.

## RESUMEN

**Objetivo.** Aplicar y valorar el enfoque bayesiano para realizar proyecciones de tasas de mortalidad por cáncer a través del ajuste de modelos edad-período-cohorta (EPC).

**Métodos.** El método de estimación bayesiano se aplica a datos de mortalidad por cáncer de vejiga en Argentina. Se adopta un esquema autorregresivo de segundo orden para la especificación a priori de los coeficientes del modelo EPC. Se comparan las estimaciones obtenidas con toda la información disponible y excluyendo los grupos de edad con tasas de mortalidad bajas, a fin de valorar el comportamiento del enfoque ante datos esparcidos. Se proyectan las tasas de mortalidad a dos períodos sucesivos a los observados.

**Resultados.** Se comprueba la robustez del método, lo cual evita excluir los grupos de edad con tasas de mortalidad nulas o bajas. Las tasas observadas caen todas dentro de las bandas de credibilidad y confirman la bondad del ajuste del modelo. Se observa una tendencia general decreciente de las tasas de mortalidad por cáncer de vejiga. Las estimaciones y proyecciones de estas tasas resultan más precisas en los grupos etarios que presentan mayor incidencia de mortalidad.

**Conclusiones.** La formulación bayesiana utilizada permite reducir la variación aleatoria entre estimaciones adyacentes al especificar que los efectos de cada escala dependen de los inmediatos anteriores. Se demuestra la capacidad del enfoque para manejar frecuencias bajas y obtener estimaciones confiables de las tasas de mortalidad, como así también proyecciones precisas sin necesidad de realizar supuestos adicionales como sucede en el ajuste clásico de un modelo EPC.

## Palabras clave

Modelos estadísticos; epidemiología y bioestadística; mortalidad; neoplasias urológicas; Argentina.

En el campo de la salud pública, es de particular interés disponer de buenas proyecciones de tasas de incidencia y mortalidad por causas específicas, ya que las mismas son de vital importancia para elaborar cualquier planificación de políticas para el sector. Esto se debe a que las tasas basadas en datos de registro

sistemático suelen publicarse con una apreciable demora y, por lo tanto, elaborar proyecciones confiables permitiría, por un lado, obtener una buena estimación de las tasas reales en el corto plazo y, en el largo plazo, contar con instrumentos significativos desde el punto de vista de la planificación.

En los últimos años, con el fin de obtener proyecciones específicas por edad de las tasas de mortalidad o incidencia de cáncer, se comenzó a utilizar con mayor frecuencia el enfoque bayesiano aplicado a los modelos edad-período-cohorta (EPC).

Estos modelos se usan frecuentemente para analizar tasas de mortalidad y morbilidad clasificadas por grupos de edad y período (1). Su estimación por métodos bayesianos permite combinar conocimientos previos con los datos observados para derivar las distribuciones de las cuales se extraen las inferencias acerca de los parámetros del modelo y funciones de dichos parámetros. Su utilización fue limitada en el pasado debido a que estas distribuciones a posteriori son difíciles o imposibles de obtener. En la actualidad, se dispone de procedimientos

<sup>1</sup> Instituto de Investigaciones Teóricas y Aplicadas de la Escuela de Estadística. Universidad Nacional de Rosario, Rosario, Argentina. La correspondencia se debe dirigir a Nora Arnesi. Correo electrónico: narnesi@fcecon.unr.edu.ar

computacionales intensivos que han facilitado su implementación mediante la aplicación de métodos aproximativos para obtenerlas. Se destaca el denominado *Markov Chain Monte Carlo* (MCMC), que combina dos métodos: el de integración Montecarlo y el muestreo según cadenas de Markov (2, 3). La formulación bayesiana supone algún tipo de suavizado de los efectos de edad, período y cohorte y es particularmente útil para predecir tasas futuras, ya que las especificaciones a priori establecidas para los efectos período y cohorte pueden ser estocásticamente extrapoladas para valores futuros (4).

Los datos de incidencia y mortalidad por cáncer suelen presentarse como cantidades estratificadas en grupos de edad de cinco años y períodos calendarios quinquenales. Es común que, para determinados tipos de cánceres, se presenten pocos casos, lo que puede dar lugar a tablas que entrecruzan grupos de edad y períodos con celdas que contienen ceros o bien con frecuencias muy bajas.

En el análisis clásico, que utiliza como método de estimación el máximo verosímil, la presencia de ceros en los datos observados puede conducir a problemas en el ajuste del modelo, así como inestabilidad en las estimaciones de los parámetros e incluso, en algunos casos, problemas de convergencia en el método de estimación utilizado. Una solución a estos problemas consiste en excluir del análisis los grupos de edades jóvenes, en los cuales hay muy pocos eventos observados. Se ha comprobado que si se excluyen ciertas celdas de la tabla de cohortes pero se utiliza el enfoque bayesiano de estimación, se obtienen estimaciones más estables que las que se logran a partir del método de máxima verosimilitud (5, 6). Más aún, dado que la presencia de ceros no significa un problema en la implementación del ajuste de modelos EPC bajo el enfoque bayesiano, muchos autores (7–11) sugieren incluir todos los datos, sin descartar aquellos que presentan frecuencias muy bajas y/o nulas.

El objetivo de este trabajo es aplicar y valorar el enfoque bayesiano en el ajuste de un modelo EPC y su utilización para realizar proyecciones de tasas de mortalidad. A fin de ilustrar su implementación y destacar sus principales características, se aplica a información registrada sobre cáncer de vejiga en Argentina.

## MATERIALES Y MÉTODOS

Los datos a analizar provienen de un diseño observacional de cortes transversales repetidos en el tiempo, ya que se dispone de información publicada por el Ministerio de Salud de la Nación Argentina en base a registros de mortalidad por causa de la Dirección de Estadística e Información en Salud (DEIS) en años sucesivos (12). En particular, se aborda el estudio de las tasas de mortalidad por cáncer de vejiga en hombres en Argentina, para lo cual se cuenta con el número de muertes por tumor maligno de vejiga en hombres y la correspondiente población en riesgo para 16 grupos de edad de cinco años cada uno (0–4 años, ..., 70–74 años) y seis períodos quinquenales (1980–1984, ..., 2005–2009). Los datos se organizan en una tabla que clasifica el número de hombres que mueren por dicha causa según el grupo de edad y el período de registro. Esta tabla se denomina de cohortes, ya que permite una tercera dimensión de lectura a través de sus diagonales, que representan las cohortes de nacimiento. Las tasas de mortalidad correspondientes a cada grupo de edad y período se obtienen relacionando el número de muertes con la correspondiente población masculina en riesgo en cada grupo. El modelo EPC aplicado a una tabla de cohortes describe la tasa de mortalidad en términos de efectos aditivos para edad, período y cohorte a través de la función de enlace logaritmo (13, 14). Esto es, para cada grupo de edad  $e$  y cada período  $p$  se dispone de información del número de muertes por alguna causa específica y el número de personas-años expuestos al riesgo de morir ( $N_{ep}$ ). Se asume que el número de muertes ( $Y_{ep}$ ) sigue una distribución Poisson, con  $E(Y_{ep}) = \mu_{ep}$  y el modelo se formaliza entonces:

$$\ln\left(\frac{\mu_{ep}}{N_{ep}}\right) = \alpha_e + \beta_p + \gamma_c \quad 0 \quad [1]$$

$$\ln(\mu_{ep}) = \alpha_e + \beta_p + \gamma_c + \ln(N_{ep}),$$

$$e: 1, 2, \dots, E; p: 1, 2, \dots, P; c: 1, 2, \dots, E + P - 1$$

Donde:  $\alpha$  es el efecto asociado al grupo etario, *edad*, al que pertenece el sujeto al momento de sufrir el evento "muerte",  $\beta$  es el efecto asociado al período en el cual

se registra esa muerte y  $\gamma$  es el efecto asociado a la cohorte a la cual pertenece dicho sujeto, es decir aproximadamente su año de nacimiento. Cabe aclarar que el subíndice  $c$  correspondiente al efecto  $\gamma$  se determina al asignar al individuo un grupo de edad y período de acuerdo a:  $c = E + p - e$ , lo cual se traduce en que el subíndice asociado a cohorte varíe entre 1 y  $(E + P - 1)$ .

La relación lineal entre los tres factores se debe a que *cohorte = período - edad* causa problemas en la identificación e interpretación de los parámetros del modelo, aunque afortunadamente no afecta las tasas estimadas ni las proyecciones de manera unívoca (15–18). Sin embargo, aún persiste el problema de la variación aleatoria que provoca inestabilidad en las estimaciones de los parámetros relacionados con edad, período y cohorte. Las proyecciones basadas en esas estimaciones, en particular de los efectos de cohortes más recientes, resultan sensibles a dicha inestabilidad.

El enfoque bayesiano aplicado al ajuste de un modelo EPC supone incorporar supuestos a priori acerca del suavizado de los parámetros asociados a edad, período y cohorte. En general, se adopta un esquema autorregresivo de segundo orden para la especificación a priori de los coeficientes del modelo como forma de suavizar los efectos en cada escala del tiempo y para extrapolar efectos de períodos y cohortes (19, 20). De acuerdo con esta especificación, cada punto, excepto los primeros dos de cada escala, se predice por extrapolación lineal a partir de sus dos predecesores inmediatos. De esta manera se evita que las estimaciones de los parámetros difieran mucho de las obtenidas para los efectos adyacentes en cada escala de tiempo. Es posible, entonces, usar el modelo autorregresivo para extrapolar los efectos de período y cohorte hacia adelante a fin de obtener las proyecciones de las tasas para períodos futuros.

Por lo tanto, se adopta la formulación bayesiana que establece:

$$\alpha_1 \sim Normal\left(0, 10^4 \frac{1}{\tau_\alpha}\right);$$

$$\alpha_2/\alpha_1 \sim Normal\left(0, 10^4 \frac{1}{\tau_\alpha}\right);$$

$$\alpha_e/\alpha_1, \dots, \alpha_{e-1} \sim Normal\left(2\alpha_{e-1} - \alpha_{e-2}, \frac{1}{\tau_\alpha}\right) \quad e > 2$$

$$\beta_1 \sim \text{Normal}\left(0, 10^4 \frac{1}{\tau_\beta}\right);$$

$$\beta_2/\beta_1 \sim \text{Normal}\left(0, 10^4 \frac{1}{t_b}\right);$$

$$\beta_p/\beta_1, \dots, \beta_{p-1} \sim \text{Normal}\left(2\beta_{p-1} - \beta_{p-2}, \frac{1}{\tau_\beta}\right) p > 2$$

$$\gamma_1 \sim \text{Normal}\left(0, 10^4 \frac{1}{\tau_\gamma}\right);$$

$$\gamma_2/\gamma_1 \sim \text{Normal}\left(0, 10^4 \frac{1}{\tau_\gamma}\right);$$

$$\gamma_c/\gamma_1, \dots, \gamma_{c-1} \sim \text{Normal}\left(2\gamma_{c-1} - \gamma_{c-2}, \frac{1}{\tau_\gamma}\right) c > 2$$

$$\tau_a \sim \Gamma(0,001, 0,001); \tau_b \sim \Gamma(0,001, 0,001); \tau_\gamma \sim \Gamma(0,001, 0,001)$$

En esta especificación del enfoque bayesiano se piensa la precisión  $\tau$ , inversa de la variancia, de cada una de las distribuciones normales como un “hiperparámetro” para el cual se especifica una distribución a priori que refleje la creencia previa relacionada con el grado de suavizado de los parámetros. Se utilizan distribuciones gamma no informativas para el “hiperparámetro” asociado a cada escala de tiempo, a fin de que estos se estimen sólo a partir de la información suministrada por los datos.

Para tener en cuenta el problema de no identificación de los parámetros del modelo se imponen restricciones a ellos, a fin de favorecer la estabilidad numérica. Dichas restricciones son: igualar a cero la tendencia lineal de los efectos asociados a edad, y provocar la suma cero de los efectos asociados a período y cohorte, restándoles sus respectivas medias.

Mediante la implementación del método MCMC en *software* BUGS (Estimación bayesiana mediante el muestreo de Gibbs) (21–24), se derivan las distribuciones a posteriori de las cuales se extraen las inferencias acerca de los parámetros del modelo y funciones de esos parámetros como, por ejemplo, las tasas específicas por edad. Se utilizan el método de cadenas múltiples y el cálculo de la estadística de Gelman-Rubin, modificada por Brooks y Gelman (25) como medios para el análisis formal de la convergencia del método. El proceso de obtención de los resultados se realiza mediante el descarte de las 5 000 iteraciones iniciales y reteniendo

el valor obtenido cada 10 iteraciones a fin de reducir la autocorrelación.

Las estimaciones puntuales se resumen por la mediana de las distribuciones marginales a posteriori de los parámetros del modelo y se calculan los intervalos de credibilidad de 95%. Las tasas estimadas se obtienen mediante la combinación de los efectos suavizados de edad, período y cohorte de acuerdo a [1]. Para ellas, también se calculan los intervalos de credibilidad.

La metodología descripta se aplica a la tabla de cohortes correspondiente a las tasas de mortalidad por cáncer de vejiga en Argentina, que identifica 16 grupos de edad, seis períodos y, por lo tanto, nueve cohortes.

Por tratarse de una patología con escasa presencia en edades jóvenes, la tabla de cohortes que se analiza tiene información dispersa, ya que algunas de sus celdas tienen frecuencias nulas o muy bajas de mortalidad. Este hecho se utiliza para ilustrar la robustez del método bayesiano. Para ello, se realizan proyecciones considerando los primeros cuatro de los seis períodos para el ajuste del modelo en dos grupos de información: la tabla completa y una tabla que excluye los grupos de edad con frecuencias nulas o muy bajas (0–4 años, ..., 30–35). Se evalúan luego las diferencias entre las tasas de mortalidad proyectadas en base al modelo para los dos períodos descartados en ambos grupos de información, completa y reducida, con las tasas realmente observadas.

A partir del modelo EPC estimado, se proyectan las tasas de mortalidad por cáncer de vejiga en hombres en Argentina para dos períodos futuros sucesivos y se calculan los intervalos de credibilidad de 95% tanto para las estimaciones como para las proyecciones.

## RESULTADOS

Un primer resultado a destacar es la comprobación de la robustez del método bayesiano ante la presencia de frecuencias nulas o bajas. En la comparación entre las tasas proyectadas y las realmente observadas para los períodos 5 (2000–2004) y 6 (2005–2009) se detectan diferencias en la precisión según se utilice toda la información disponible o sólo una parte de ella.

A modo de ejemplo, en la figura 1 se presenta una comparación de las tasas proyectadas basadas en el conjunto de

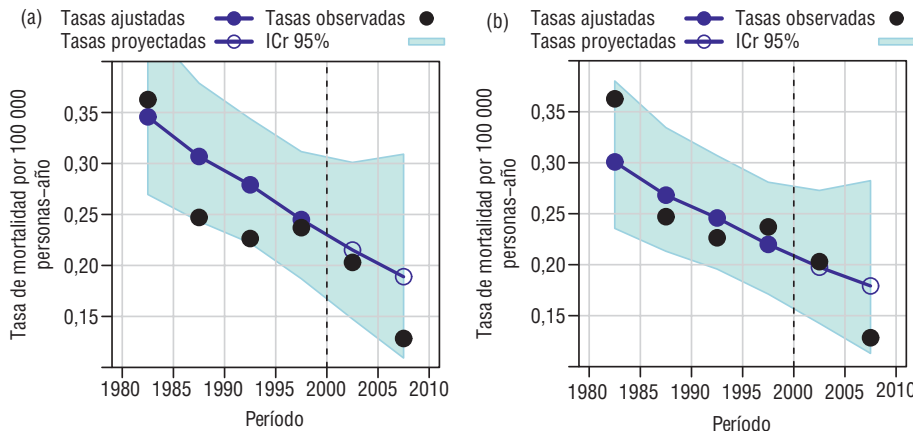
datos completo (panel a) y en el reducido (panel b), para un grupo de edad particular: 35–39 años. En dicha figura es posible observar una mayor amplitud en el intervalo de credibilidad de 95% cuando las proyecciones se realizan en base al conjunto de datos reducido. Como una forma de resumir este comportamiento a través de todas las edades que están tanto en el conjunto completo como en el reducido, se calculan medidas descriptivas de los sesgos, en valor absoluto, entre la tasa observada y la proyectada. Así, para el período 5 (2000–2004), la media de los sesgos en el grupo de información reducida resultó igual a 0,536 mientras que considerando la información completa alcanza un valor de 0,508. Para el período 6 (2005–2009), la media de los sesgos resulta de 1,503 y 1,435 para cada uno de los grupos, respectivamente. También la mediana de los sesgos resulta levemente inferior cuando se incluyen todos los grupos de edad, lo que indica que disponer de más información, incluso la de aquellos grupos de edad con frecuencias muy bajas, mejora las proyecciones.

Una vez comprobada la utilidad de considerar toda la información disponible, es decir los seis quinquenios que abarcan desde 1980 a 2009, para los 16 grupos de edad, se ajusta el modelo EPC bajo el enfoque bayesiano propuesto y luego se proyectan las tasas de mortalidad para dos períodos a futuro: período 7 (2010–2014) y período 8 (2015–2019), es decir para un período total de 10 años.

Un resultado interesante de señalar tiene que ver con los valores obtenidos como estimación a posteriori de los parámetros de precisión,  $\tau$ , de los tres factores incluidos en el modelo los cuales resultaron 11,97 para edad, 461,50 para período y 1 402,00 para cohorte. Estos valores sugieren que los efectos sobre los ejes período y cohorte están más suavizados que sobre el eje de la edad. En otras palabras, significa que la edad aparece como un factor considerablemente más influyente que el período de registro o la cohorte de nacimiento sobre las tasas de mortalidad por cáncer de vejiga en hombres en Argentina.

En la figura 2 se presentan las tasas ajustadas y las proyectadas con el intervalo de credibilidad de 95% en tres grupos de edad elegidos a modo de ejemplo: 35–39, 45–49 y 74–79, junto con las tasas observadas (paneles a, b y c). La elección de estos grupos de edad obedece a la intención de mostrar la gran variación existente

**FIGURA 1. Tasas de mortalidad observadas, ajustadas y proyectadas con intervalo de credibilidad de 95% (ICr 95%) para cáncer de vejiga en hombres de 35 a 39 años en Argentina, en base al conjunto de datos reducido (panel a) y conjunto de datos completo (panel b)**



Fuente: elaboración propia en base a los resultados obtenidos en el presente trabajo.

entre los valores de las tasas de mortalidad a medida que la población envejece.

En los tres grupos de edad se observa una tendencia decreciente en las tasas a través de los períodos incluso en las proyectadas para los dos períodos futuros. Las tasas observadas caen todas dentro de las bandas de credibilidad, lo que confirma la bondad del ajuste del modelo. Por otro lado, la amplitud de las bandas de credibilidad en relación a las tasas ajustadas indica que las estimaciones resultan más precisas en el grupo de edad de 74–79 años, esto se explica por la mayor incidencia de mortalidad por cáncer de vejiga en los grupos de edades más avanzadas. Por el contrario, en grupos de edades más jóvenes con menor incidencia, se obtienen estimaciones menos precisas. Además, es posible evaluar a través de la amplitud de dichas bandas la incertidumbre ligada a las proyecciones para períodos futuros. En este sentido, se destaca la mayor amplitud de los intervalos a medida que las proyecciones se alejan más del rango de datos observados en los tres grupos de edad ejemplificados.

## DISCUSIÓN

En este trabajo se muestran las principales características de la metodología bayesiana aplicada a un análisis EPC en un caso particular: las tasas de mortalidad por cáncer de vejiga en hombres en Argentina.

Argentina se encuentra dentro del rango de países con incidencia de cáncer media a alta y, además, es la segunda causa de muerte. En cuanto al cáncer de

vejiga, es el segundo tumor urológico en frecuencia, suele ser agresivo y ocurrir en personas de edad avanzada. Sin embargo, las tasas de mortalidad por cáncer de vejiga en hombres en el mundo se han mantenido bastante estables (26, 27).

Los resultados encontrados para Argentina confirman que el riesgo de muerte por cáncer de vejiga aumenta con la edad. La tendencia decreciente encontrada para los últimos períodos como así también las proyecciones para períodos futuros podría explicarse en gran medida por los importantes avances en los últimos tiempos tanto en la imagenología como en tratamientos quimioterápicos y de radioterapia (28). Por ejemplo, existe hoy en día una mayor posibilidad de realizar mejores estudios por imágenes que permiten realizar prácticas diagnósticas y tratamiento endoscópico en un mismo tiempo. Hay opciones que permiten realizar quimioterapia posoperatoria inmediata. No hay duda de que estos avances permiten mejoras en el diagnóstico y tratamiento pero además, como en otras patologías, la detección temprana es esencial para evitar la progresión de la enfermedad y sus complicaciones.

Al mismo tiempo, el análisis de las tasas de mortalidad por cáncer de vejiga permite mostrar a nivel empírico las propiedades de la metodología bayesiana propuesta. Esta patología se eligió precisamente por presentar escasa frecuencia en los grupos de edades más jóvenes y, por lo tanto, posibilitó examinar el impacto de la existencia de datos desbalanceados en la inferencia. Luego de com-

probar la capacidad del método para manejar ceros o frecuencias bajas, se encontró que las proyecciones a 10 años basadas en el conjunto de datos completo, sin descartar los grupos de edad con baja frecuencia, evidencian un leve aumento en la precisión cuando se realizan en grupos de edad con incidencia más alta.

Estos resultados analíticos pueden ser considerados como más confiables que otros que se pudieran derivar de métodos más sensibles a la falta de un tamaño de muestra adecuado, como lo es el máximo verosímil, con buenas propiedades pero de carácter asintótico. Por ejemplo, se ha demostrado que en estudios con menos de 10 períodos, el método del máximo verosímil no provee estimaciones precisas (29).

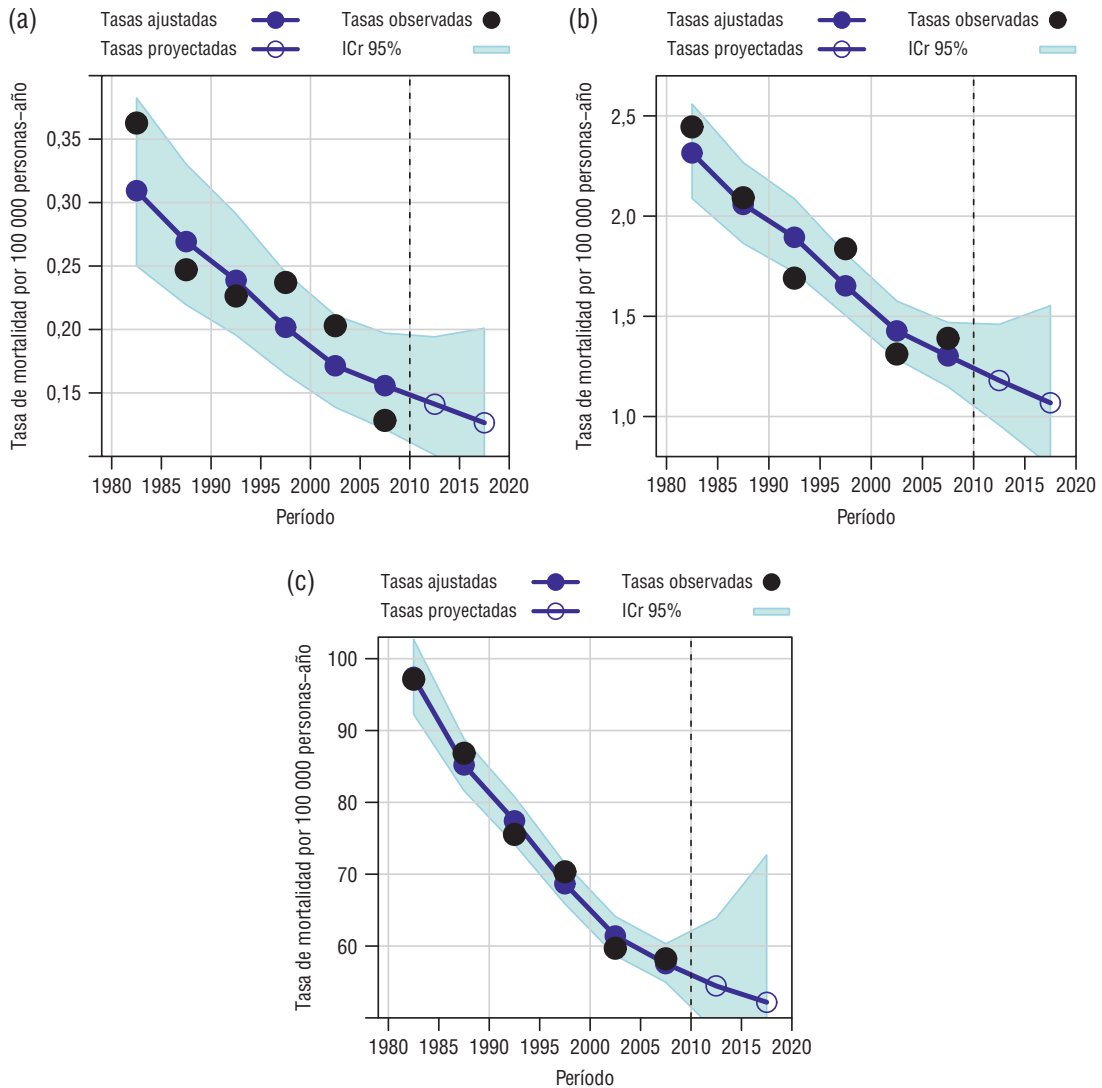
La especificación del enfoque bayesiano elegida establece “priors” autorregresivas de segundo orden para los niveles de cada escala del tiempo en el modelo EPC. De esta manera, cada estimación cobra fuerza de los datos relevantes de dos períodos anteriores y evita así una gran variación con respecto a las estimaciones adyacentes (30).

El modelo EPC así especificado permite realizar proyecciones para tiempos futuros mediante la extrapolación de efectos de cohortes y períodos (4, 31). El esquema bayesiano predice tasas futuras haciendo fluir la información contenida en el pasado hacia las estimaciones de los parámetros del futuro; de este modo fortalece la estimación al pedir prestada información de las bandas vecinas. Por el contrario, bajo el enfoque clásico EPC, raras veces se realizan proyecciones debido a que es necesario hacer supuestos paramétricos cuestionables (4). Además, si bien durante mucho tiempo el enfoque bayesiano ha ofrecido una alternativa válida, era de difícil aplicación por los problemas computacionales que supone. Esta dificultad parece saldada en la actualidad.

Se pueden hacer recomendaciones sobre la base de los resultados encontrados: es posible usar toda la información disponible y, además, obtener en forma sencilla buenas proyecciones bayesianas en un modelo EPC, siempre y cuando la etiología sea similar en todas las edades.

Los hallazgos están orientados a provocar que el analista privilegie el esquema bayesiano en un análisis EPC aplicado a un problema de salud. En particular, el cáncer ha sido y continúa siendo una preocupación, y cualquier política al respecto depende del conocimiento de las tasas de mortalidad y del monitoreo de las tendencias. Los mode-

**FIGURA 2. Tasas de mortalidad con intervalo de credibilidad de 95% (ICr 95%) por cáncer de vejiga en hombres en Argentina para los grupos de edad 35–39 años (panel a); 45–49 años (panel b) y 75–79 años (panel c)**



**Fuente:** elaboración propia en base a los resultados obtenidos en el presente trabajo.

los EPC pueden constituir un primer paso para análisis posteriores y su estimación mediante el enfoque bayesiano resulta un buen método para realizar proyecciones las cuales juegan un rol clave en la toma de decisiones gubernamentales, a fin de prevenir incrementos en la morbilidad y mortalidad.

A través de una aplicación, la intención principal de este trabajo es familiarizar a los especialistas en salud pública y epidemiólogos con los métodos bayesianos, en el convencimiento de que se trata de una metodología, aún de escasa utilización, potente para brindar información confiable a los tomadores de decisiones.

**Conflicto de intereses.** Ninguno declarado por los autores.

**Declaración.** Las opiniones expresadas en este manuscrito son responsabilidad de los autores y no reflejan necesariamente los criterios ni la política de la RPSP/PAJPH y/o de la OPS.

**REFERENCIAS**

1. Clayton D, Schifflers E. Models for temporal variation in cancer rates I: age-period-cohort models. *Statist Med.* 1993;6:449–67.
2. Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences. *Statist Sci.* 1992;7:457–72.
3. Gilks WR, Richardson S, Spiegelhalter DJ. *Markov Chain Monte Carlo in Practice.* Interdisciplinary Statistics. Oxford: Chapman&Hall; 1996.
4. Bray I. Application of Markov chain Monte Carlo methods to projecting cancer incidence and mortality. *Appl Stat.* 2002;51:151–64.
5. Raifu AO, Arbyn M. Description of cervical cancer mortality in Belgium using Bayesian age-period-cohort models. *Arch Public Health.* 2009;67:100–15.

6. Arnesi N, Prunello M, Marfetan Molina D, Hachuel L. El problema de datos esparcidos en los modelos edad-período-cohorte: enfoque bayesiano versus clásico. Libro de Resúmenes del XV Congreso y XXXIII Reunión Anual de la Sociedad de Biología de Rosario. 2013. Disponible en: <http://www.sbr.org.ar> Acceso en marzo de 2014.
7. Baker A, Bray I. Bayesian Projections: what are the effects of excluding data from younger age groups? *Am J Epidemiol*. 2005;162:798–805.
8. Knorr-Held R, Rainer E. Projections of lung cancer mortality in West Germany: a case study in Bayesian prediction. *Biostatistics*. 2001;2:109–29.
9. Verdecchia A, De Angelis G, Capocaccia R. Estimation and projections of cancer prevalence from cancer registry data. *Stat Med*. 2002;21:3511–26.
10. Brennan P, Bray I. Recent trends and future directions for lung cancer mortality in Europa. *Br J Cancer*. 2002;87:43–8.
11. Bray I, Brennan P, Boffetta P. Recent trends and future projections of lymphoid neoplasms – A Bayesian age-period-cohort analysis. *Cancer Causes Control*. 2001; 12:813–20.
12. Ministerio de Salud de la República Argentina. Dirección Estadística e Información en Salud. Anuarios e informe especial: evolución de la mortalidad según sus causas. Disponible en: <http://www.deis.gov.ar> Acceso en julio de 2013.
13. Agresti A. *Categorical Data Analysis*. 3<sup>o</sup> ed. Wiley, 2013.
14. Dobson A, Barnett G. *An Introduction to Generalized Linear Models*. 3<sup>o</sup> ed. Nueva York: CRC Press; 2008.
15. Clayton D, Schifflers, E. Models for temporal variation in cancer rates II: age-period-cohort models. *Statist Med*. 1987;6:469–81.
16. Holford, TR. An alternative approach to statistical age-period-cohort analysis. *J Chron Dis*. 1985;38:831–6.
17. Holford, TR. Understanding the effects of age, period and cohort on incidence and mortality rates. *A Rev Publ Hlth*. 1991;12: 425–57.
18. Holford, TR. Analysing the temporal effects of age, period and cohort. *Statist Meth Med Res*. 1992;1:317–37.
19. Breslow N, Clayton DG. Approximate inference in generalized linear mixed models. *J Am Statist Ass*. 1993;88:9–25.
20. Berzuini C, Clayton D. Bayesian analysis of survival on multiple time scales. *Statist Med*. 1994;13:823–38.
21. Ntzoufras I. *Bayesian Modeling Using WinBUGS*. Wiley Series in Computational Statistics. Wiley, 2009.
22. Robert C, Casella J. *Use R! Introducing Monte Carlo Methods with R*. Nueva York: Springer; 2010.
23. Marin JM, Robert CM. *Bayesian Core: A Practical Approach to Computational Bayesian Statistics*. Nueva York: Springer; 2007.
24. R Core Team. *A language and environment for statistical computing*. Viena: R Foundation for Statistical Computing; 2013. Disponible en <http://www.r-project.org/> Acceso 2015.
25. Brooks S, Gelman A. *General Methods for Monitoring Convergence of Iterative Simulations*. 1997. Disponible en: <https://stat.duke.edu/~scs/Courses/Stat376/Papers/ConvergeDiagnostics/BrooksGelman.pdf> Acceso en marzo de 2014.
26. Ministerio de Salud de la Nación Argentina. Instituto Nacional del Cáncer. Análisis de la situación del cáncer en Argentina. Disponible en: <http://www.msal.gov.ar/inc/index.php/acerca-del-cancer/estadistica> Acceso en septiembre de 2014.
27. American Cancer Society. 2014. Cáncer de vejiga. Disponible en: <http://www.cancer.org/acs/groups/cid/documents/webcontent/002280.pdf> Acceso en diciembre de 2014.
28. Federación Argentina de Urología. Sociedad Argentina de Urología. Consenso Urológico Nacional sobre pautas para el Diagnóstico y Tratamiento del Cáncer de Vejiga. 2013. Disponible en: [www.sau-net.org/publicaciones/lineamientos-diagnostico-tratamiento/lineamientos-tumores](http://www.sau-net.org/publicaciones/lineamientos-diagnostico-tratamiento/lineamientos-tumores) Acceso en junio de 2014.
29. Yang Y. *Bayesian Inference for Hierarchical Age-Period-Cohort Models of Repeated Cross-Section Survey Data*. *Sociol Methodol*. 2006;36(1):39–74.
30. Riebler A, Held L. The analysis of heterogeneous time trends in multivariate age-period-cohort models. *Biostatistics*. 2009; 11:57–69.
31. Schmid VJ, Held L. Bayesian age-period-cohort modeling and prediction-BAMP. *J Stat Software*. 2007;21(8):1–15.

Manuscrito recibido el 2 de noviembre de 2014. Aceptado para publicación, tras revisión, el 15 de junio de 2015.

## ABSTRACT

### Use of a Bayesian approach to rate estimation and projection

**Objective.** Apply and assess a Bayesian approach to projecting cancer mortality rates by fitting age-period-cohort (APC) models.

**Methods.** The Bayesian estimation method was applied to bladder cancer mortality data in Argentina. A second-order autoregressive model was adopted for a priori specification of APC model coefficients. The estimates obtained were compared with all available information and excluding age groups with low mortality, to assess behavior of the approach in light of scattered data. Mortality was projected for two successive periods following the ones observed.

**Results.** Robustness of the method was verified, which avoids excluding age groups with null or low mortality. Observed rates all fall within the credibility bands and confirm the model's goodness of fit. An overall downward trend in bladder cancer mortality was observed. Estimates and projections of these rates are more precise in age groups that have greater incidence of mortality.

**Conclusions.** The Bayesian formulation used herein makes it possible to reduce random variation between adjacent estimates by specifying that the effects of each scale depend on the immediately preceding ones. It was demonstrated that the approach has the capacity to handle low frequencies and obtain reliable mortality estimates, as well as precise projections, without the need for making additional assumptions, as happens in classical APC model fitting.

## Key words

Models, statistical; epidemiology and biostatistics; mortality; urologic neoplasms; Argentina.