Lung cancer mortality trends in Chile and six-year projections using Bayesian dynamic linear models

Mortalidad por cáncer pulmonar en Chile: tendencia y proyección a seis años con modelos lineales dinámicos bayesianos

Mortalidade por câncer de pulmão no Chile: tendência e projeção de seis anos com modelos lineares dinâmicos bayesianos

Francisco Torres-Avilés †
Tomás Moraga [1]
Loreto Núñez [2,3]
Gloria Icaza [3,4]

[1] *Departamento de Matemática y Ciencia de la Computación, Universidad de Santiago de Chile, Santiago, Chile.*
[2] *Departamento de Salud Pública, Universidad de Talca, Talca, Chile.*
[3] *Programa de Investigación Interdisciplinaria de Excelencia en Envejecimiento Saludable, Universidad de Taca, Talca, Chile.*
[4] *Instituto de Matemática y Física, Universidad de Talca, Talca, Chile.*
† *Deceased.*

Correspondence
G. Icaza
Instituto de Matemática y Física, Universidad de Talca.
Casilla 721, Talca
3460-000, Chile.
gicaza@utalca.cl

Abstract

The objectives were to analyze lung cancer mortality trends in Chile from 1990 to 2009, and to project the rates six years forward. Lung cancer mortality data were obtained from the Chilean Ministry of Health. To obtain mortality rates, population projections were used, based on the 2002 National Census. Rates were adjusted using the world standard population as reference. Bayesian dynamic linear models were fitted to estimate trends from 1990 to 2009 and to obtain projections for 2010-2015. During the period under study, there was a 19.9% reduction in the lung cancer mortality rate in men. In women, there was increase of 28.4%. The second-order model showed a better fit for men, and the first-order model a better fit for women. Between 2010 and 2015 the downward trend continued in men, while a trend to stabilization was projected for lung cancer mortality in women in Chile. This analytical approach could be useful implement surveillance systems for chronic non-communicable disease and to evaluate preventive strategies.

Lung Neoplasms; Bayes Theorem; Mortality Rate; Health Surveillance

Resumen

El objetivo fue analizar la tendencia de la tasa de mortalidad por cáncer de pulmón en Chile, durante el periodo 1990-2009 y proyectar estas tasas a seis años. La información de mortalidad fue obtenida del Ministerio de Salud de Chile. Para calcular las tasas se utilizaron las proyecciones de población según el Censo de 2002. Las tasas se estandarizaron usando la población mundial como referencia. Se ajustaron modelos lineales dinámicos bayesianos para estimar la tendencia entre 1990-2009 y proyectar el periodo 2010-2015. Durante el periodo se observa una reducción del 19,9% de la tasa de mortalidad en hombres, mientras que en mujeres, la tendencia es creciente con aumento de 28,4%. El modelo de segundo orden entregó un mejor ajuste en hombres y el de primer orden en mujeres. Entre 2010 y 2015, se mantiene la tendencia decreciente en hombres, en cambio se proyecta una estabilización en la tendencia de mortalidad por cáncer pulmonar en mujeres en Chile. Este tipo de análisis es útil para implementar sistemas de vigilancia epidemiológica y evaluar estrategias.

Neoplasias Pulmonares; Teorema de Bayes; Tasa de Mortalidad; Vigilancia Sanitaria

Introduction

Lung cancer accounts for 1.6 million deaths per year according to the World Health Organization (WHO) [1]. In Chile, cancer of the trachea, bronchi, and lungs (hereinafter "lung cancer") is the second cause of cancer mortality, following gastric cancer (Departamento de Estadísticas e Información en Salud, Ministerio de Salud. Mortalidad. http://www.deis.cl/?p=51, accessed on 23/Jul/2013).

In developed countries, lung cancer incidence and mortality rates are declining in men and stabilizing in women thanks to changes in smoking prevalence. However, according to estimates in developing countries, incidence and mortality continue to increase due to endemic smoking prevalence [2].

In Chile, the lung cancer mortality rate showed an upward trend in females and a slightly downward trend in males from 2001 to 2008 [3]. The increase in the female population can be explained by the late adoption of smoking by women, due to sociocultural issues and exploitation of this characteristic by the tobacco industry [4].

Time trend analysis of indicators like mortality is useful for monitoring a country's health status and the impact of health interventions [5]. Together with time trend analyses, it is useful to estimate the magnitude of the disease in the future, thus allowing optimization of resource allocation, services planning, and public policy-making [6].

The current study aimed to analyze trends in the lung cancer mortality rate in men and women in Chile from 1990 to 2009 and to conduct projections for the next six years, using Bayesian dynamic linear models.

Methods

Data source

Data on lung cancer mortality [International Classification of Diseases, 9th revision (ICD-9): 162 from the year 1990 to 1996 and C33-C34 from the 10th revision (ICD-10): for the years 1997 to 2009], data were obtained from the available databases of the Department of Health Statistics and Information (DEIS) of the Ministry of Health of Chile, 1990-2009 (http://www.deis.cl/?p=51, accessed on 23/Jul/2013). Population data were obtained from projections based on the *2002 Census* conducted by the Chilean National Institute of Statistics and the Latin American and Caribbean Demographic Center [7].

Time trend models

Standardized annual mortality rates were calculated by five-year age groups from 1990 to 2009 for men and women. The standardization method was direct adjustment, using the world standard as the reference population [8].

The study used Bayesian dynamic linear models (DLM), where Y_t represents the log of the mortality rate in time t and the model can be represented by the following system of equations:

$$Y_t = F_t' \Theta_t + \varepsilon_t$$

$$\Theta_t = G_t \Theta_{t-1} + \omega_t$$

Where, F_t is a vector of order P x 1 which is formed by co-variables, Θ_t is the vector of the model's unknown parameters p, G_t is a p x p order matrix that describes the trend of the parameters contained in Θ_t over time, ε_t and ω_t represent random errors that are assumed to have typically normal distribution with mean 0, and variance-covariance matrices (depending on the structure of G_t) V_t and W_t respectively.

For the study's analyses, we specifically used two structures for F_t and G_t.

First-order model: $F_t = 1$, $G_t = 1$ and $\Theta_t = \Theta_{1t}$.

Second-order model: $F_t = [1 \ 0]$, $G_t = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $F_t = [1 \ 0]$ $\Theta_t = [\Theta_{1t} \ \ \Theta_{2t}]$.

What distinguishes DLM from ordinary time-series models is the specification of the autocorrelation order for the model's structural parameters, making them more flexible. If there is a degree of autocorrelation, the second-order model is the most adequate.

Estimation of Bayesian DLM requires a stepwise filtering and smoothing process, in addition to defining a priori distributions for the unknown parameters that intervene in the models. When using the Bayesian paradigm in the estimation process, it is necessary to define knowledge a priori through distributions for the unknown parameters that intervene in the models. Thus, one assumes that the variances are independent and follow non-informative inverse gamma distributions. Additionally, one includes an a priori distribution for the series' baseline structural parameter, Θ_0. The literature suggests specifying a normal distribution for the target parameter. Theoretical details on such models can be found in West & Harrison [9] and Gamerman & Lopes [10].

Computational implementation

Adjustment of these models used the code's implementation in Winbugs 1.4.3 (http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/contents.shtml). Nevertheless, it is possible to use the dlm library from the R software package (http://www.r-project.org/), whose functions dlmModPoly and dlmModSeas are designed to estimate dynamic linear models [11]. To perform inference, chains of 60,000 iterations were generated, excluding the first 40,000 to eliminate the influence of baseline values and autocorrelation, which would allow ensuring the parameters' convergence in the MCMC strategy (Markov chain Monte Carlo stochastic simulation) used. Having estimated the models' parameters, one evaluates their goodness of fit using the deviance information criterion (DIC), the effective number of parameters (pD) [12], the magnitude of the discrepancy observed *a posteriori* between the data expected by the models and the observed data (SCEp), and the model's predictive capacity based on the prediction error.

The model's projection was also compared to the data for 2010-2012, recently published by DEIS (http://www.deis.cl/?p=51, accessed on 23/Jul/2013). The study provides results of the best model, and based on which, projections of lung cancer mortality rates for the next six years in men and women in Chile, with the respective credible intervals.

Results

Time trend analysis of standardized lung cancer mortality rates in men and women in Chile from 1990 to 2009 showed a reduction of 19.9% in men, while in women there was an upward trend, namely a 28.4% increase (Figure 1).

The second-order DLM was the model with the best fit for the trend of standardized mortality rates in men, according to Bayesian goodness of fit criteria. Meanwhile, in women the first-order DLM showed the best fit (Table 1). Mortality data for the years 2010 to 2012 were used to obtain age-adjusted rates for both males and females, the value of which is contained in both models' confidence intervals (Table 2). Meanwhile, the differences between the expected and observed mortality rates were less than one (Table 2).

Based on each model, the study provided the six-year projections and credible intervals (Figures 2 and 3). In men, the downward trend observed in 1990-2009 was maintained. In women, the six-year projection showed stabilization of the lung cancer mortality rates, thereby interrupting the upward trend observed in the period under study.

Figure 1

Age-adjusted lung cancer mortality rates by sex. Chile 1990-2009.

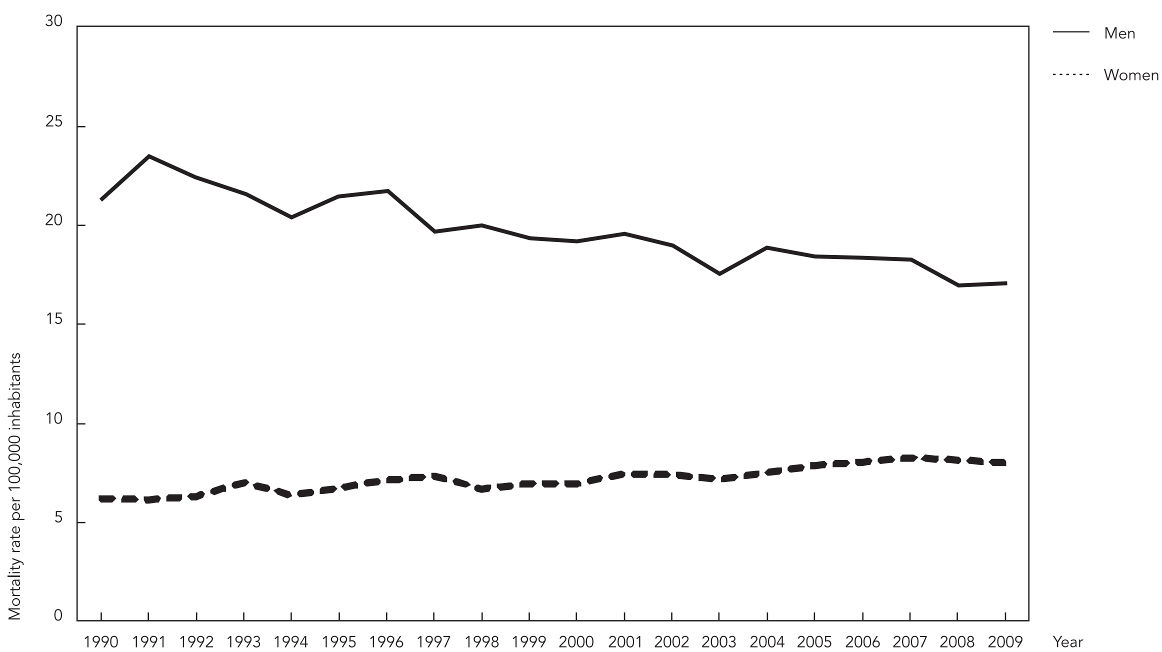


Table 1

Selection criteria used in models for lung cancer mortality rates, Chile.

Sex	Model	pD	DIC	EPD	SCEp
Men	1st-order DLM	11.945	-71.218	0.0433	0.0193
	2nd-order DLM	14.093	-69.606	0.0427	0.0192
Women	1st-order DLM	12.580	-71.670	0.0412	0.0184
	2nd-order DLM	14.139	-70.036	0.0419	0.0187

DIC: deviance information criterion; DLM: dynamic linear models; EPD: prediction error; pD: number of parameters (pD); SCEp: the magnitude of the discrepancy between the observed expected data by the *posteriori* models.

Table 2

Projected lung cancer mortality rates (2010-2012), Chile.

Sex	Year	Adjusted observed rate	Projected rate (95% credible interval)	Absolute difference
Men (2nd-order model)	2010	16.76	16.69 (14.18-19.51)	0.07
	2011	17.23	16.43 (12.72-21.00)	0.80
	2012	16.27	16.22 (11.14-22.88)	0.05
Women (1st-order model)	2010	7.74	8.01 (7.13-8.99)	0.27
	2011	8.41	8.02 (6.97-9.22)	0.39
	2012	8.75	8.03 (6.82-9.40)	0.72

Discussion

Chile is considered to have world-standard vital statistics [13] thanks to a systematic effort by the DEIS under the Ministry of Health of Chile in terms of coding and processing death certificates. Nevertheless, on-going efforts are needed to reduce the differences observed between urban and rural areas, counties, males and females, and age groups [14].

One of the methods current employed to analyze trends in mortality rates is joinpoint regression [15], useful for identifying changes in trends over time. The currently study used Bayesian DLM, which have been applied to a wide range of situations due to their predictive power, since the models' parameters are updated with each new observation, allowing them to adapt to the series' evolution and capturing changes in their behavior. Bayesian DLM are thus flexible for detecting change points and useful for forecasting [16].

The downward trend in lung cancer in men has been described in developed countries, as has the upward trend in women, reflecting smoking prevalence at the population level [2]. In the case of women, a plateau effect has been seen following the upward trend, as described in the current study.

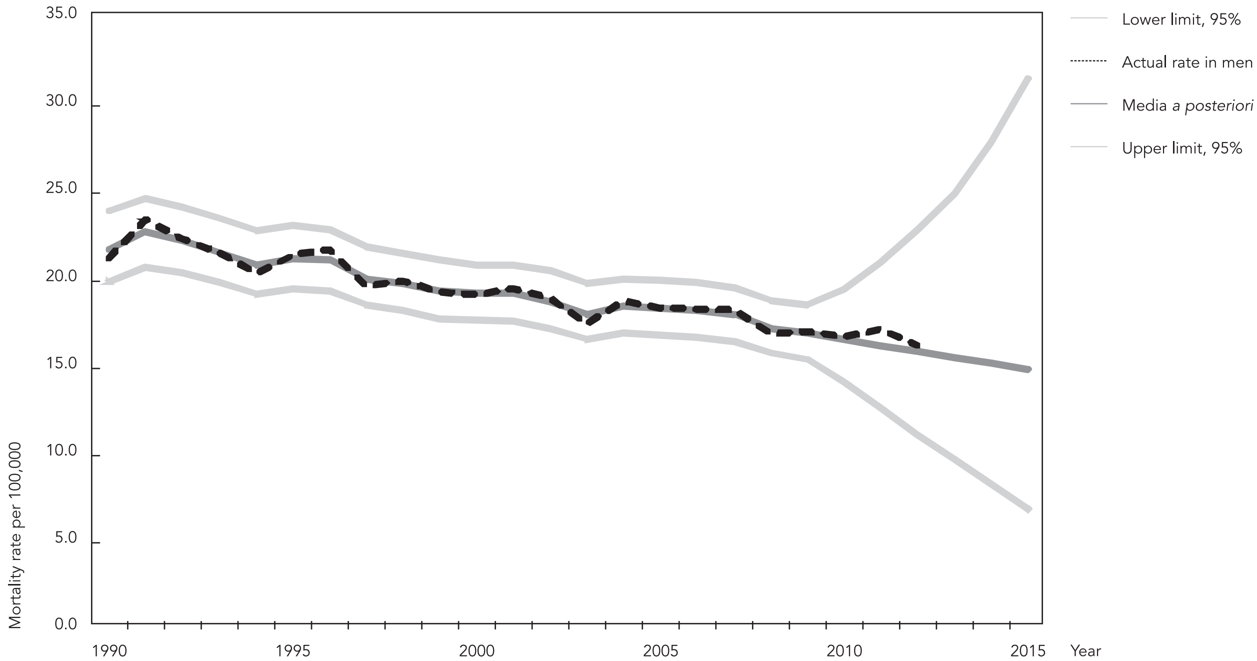
In Chile, the downward trend in men can be attributed not only to changes in smoking prevalence, but also to the longstanding exposure to arsenic in drinking water in the northern region of the country, followed by the installation of the first water treatment plants in 1971 [17].

One of study's limitations is that these rates are a summary measure for Chile as a whole and do not consider geographic differences, found in the relative risk of this disease due to the presence of arsenic in the north of the country [3]. Spatiotemporal studies are planned to take this phenomenon into account.

Another limitation is that the projections fail to consider changes in incidence of the disease or the incorporation of preventive strategies, so the results must be interpreted in the context of status quo. Importantly, Chile has adopted a series of changes in its anti-tobacco legislation which are expected to contribute to the decrease

Figure 2

Six-year projections according to second-order model and standardized lung cancer mortality rates in men. Chile 1990-2015.

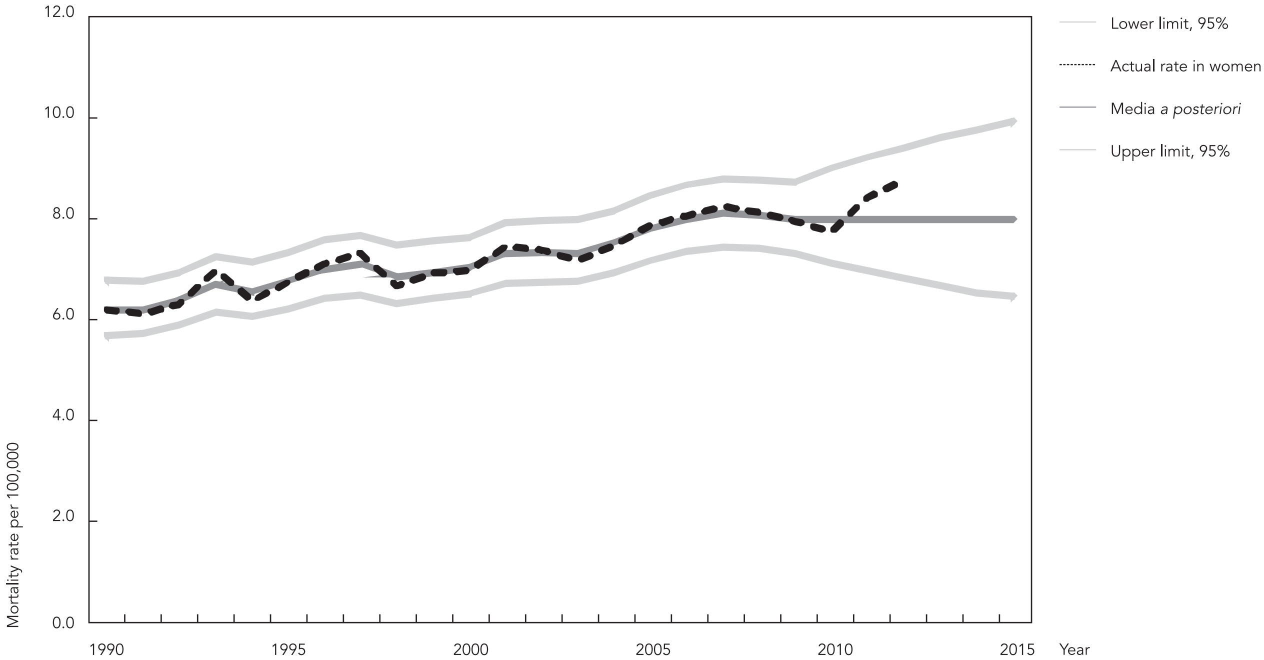


in mortality. Meanwhile, according to data from Chile's National Health Surveys, smoking prevalence in individuals more than 15 years of age decreased in males from 48.3% in 2003 to 44.2% in 2010, while in females it increased from 36.8% to 37.1% in the same period [18,19]. An aggravating factor is that Chilean schoolchildren (13 to 15 years of age) have the highest smoking rates in the Americas for their age group [20]. At the local level, studies on the general adult population in Santiago that evaluate the national trend reported 47% smoking prevalence in men and 27% in women in 1971. In 1984, the prevalence rates were 44% and 39%, respectively [21]. Finally, a study in 2003 found smoking prevalence rates of 46.5% in men and 39.4% in women [22].

The current study used a methodology that allows modeling lung cancer mortality rates and producing projections. The methodology can be replicated to forecast mortality rates from other diseases in the population and make adjustments in public policies and implement surveillance systems for chronic non-communicable diseases and evaluate appropriate preventive strategies.

Figure 3

Six-year projections according to first-order model and standardized lung cancer mortality rates in women. Chile 1990-2015.



Resumo

O objetivo foi analisar a tendência da taxa de mortalidade por câncer de pulmão no Chile, durante 1990-2009 e projetar estas taxas em seis anos. As informações sobre a mortalidade foram obtidas no Ministério da Saúde do Chile. Para calcular as taxas foram utilizadas projeções da população de acordo com o Censo de 2002. As taxas foram padronizadas utilizando-se a população mundial como referência. Modelos lineares dinâmicos bayesianos foram ajustados para estimar a tendência entre 1990-2009 e projetar o período 2010-2015. Durante o período, observa-se uma redução da taxa de mortalidade de 19,9% nos homens, entretanto nas mulheres a tendência é de crescimento, com o aumento de 28,4%. O modelo de segunda ordem deu um melhor ajuste para os homens e mulheres de primeira ordem. Entre 2010 e 2015, a tendência descendente é mantida entre os homens, no entanto, projetando uma tendência de estabilização da mortalidade por câncer de pulmão em mulheres no Chile. Este tipo de análise pode ser útil para implementar sistemas de vigilância epidemiológica e avaliar estratégias.

Neoplasias Pulmonares; Teorema de Bayes; Coeficiente de Mortalidade; Vigilância Sanitária

Contributors

F. Torres-Avilés participated in the study's conceptualization and design, adjustment of the models, data analysis and interpretation, writing of the article, and approval of the final version for publication. T. Moraga contributed on the data collection, adjustment of the models', data analysis and interpretation, writing of the article, and approval of the final version for publication. L. Núñez participated in the study's conceptualization, data analysis and interpretation, writing of the article, and approval of the final version for publication. G. Icaza colaborated on the study's conceptualization, data collection, analysis, and interpretation, writing of the article, and approval of the final version for publication.

Acknowledgments

The authors wish to thank Prof. Abel Valdebenito S. for validating the fit in the models used in the current study. To Fondo Nacional de Desarrollo Científico y Tecnológico (Fondecyt) for the financial support to Francisco Torres-Avilés (Fondecyt Iniciación 11110119).

References

1. Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, et al. GLOBOCAN 2012 v1.0, cancer incidence and mortality worldwide. Lyon: International Agency for Research on Cancer; 2013. (IARC CancerBase, 11).
2. Youlden DR, Cramb SM, Baade PD. The international epidemiology of lung cancer. J Thorac Oncol 2008; 3:818-31.
3. Icaza G, Núñez L, Torres-Avilés F, Díaz N, Villarroel JE, Soto A, et al. Atlas de mortalidad en Chile, 2001-2008. Talca: Editorial Universidad de Talca; 2013.
4. Nerín I. El tabaquismo en la mujer: una atracción fatal. Arch Bronconeumol 2005; 41:360-2.
5. Puig X, Ginebra J, Gispert R. Análisis de la evolución temporal de la mortalidad mediante modelos lineales generalizados. Gac Sanit 2005; 19:481-5.
6. Mistry M, Parkin DM, Ahmad AS, Sasieni P. Cancer incidence in the United Kingdom: projections to the year 2030. Br J Cancer 2011; 105:1795-803.
7. Instituto Nacional de Estadística; Centro Latinoamericano y Caribeño de Demografía. Chile: proyecciones y estimaciones de población. Total país 1950-2050. Santiago: Instituto Nacional de Estadística/Centro Latinoamericano y Caribeño de Demografía; 2004.
8. Doll R, Payne P, Waterhouse JAH. Cancer incidence in five continents. v. I. Geneva: Union Internationale Contre le Cancer; 1966.
9. West M, Harrison J. Bayesian forecasting and dynamic models. 2nd Ed. New York: Springer; 1999.
10. Gamerman D, Lopes, H. Markov chain Monte Carlo: stochastic simulation for Bayesian inference. 2nd Ed. Boca Raton: Chapman & Hall/CRC Press; 2006.
11. Petris G. An R package for dynamic linear models. J Stat Softw 2010; 36:1-16.
12. Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. Bayesian measures of model complexity and fit. J R Stat Soc B 2002; 64:583-639.
13. Mahapatra P, Shibuya K, Lopez AD, Coullare F, Notzon FC, Rao C, et al. Civil registration systems and vital statistics: successes and missed opportunities. Lancet 2007; 370:1653-63.
14. Núñez L, Icaza G. Calidad de las estadísticas de mortalidad en Chile, 1997-2003. Rev Méd Chile 2006; 134:1191-6.
15. Kim H-J, Fay M, Feuer E, Midthune D. Permutation tests for joinpoint regression with applications to cancer rates. Stat Med 2000; 19:335-51.
16. Fonseca F, Soares A, Telles P, Williamson D. Dynamic linear model and SARIMA: a comparison of their forecasting performance in epidemiology. Stat Med 2001; 20:3051-69.
17. Steinmaus CM, Ferreccio C, Acevedo J, Yuan Y, Cortes S, Marshall G, et al. Drinking water arsenic in northern Chile: high cancer risks 40 years after exposure cessation. Cancer Epidemiol Biomarkers Prev 2013; 22:623-30.
18. Ministerio de Salud. Resultados: I Encuesta de Salud, Chile 2003. http://epi.minsal.cl/epi/html/invest/ens/informefinalens.pdf (accessed on 11/May/2013).
19. Ministerio de Salud. Encuesta Nacional de Salud 2009-2010. http://www.minsal.gob.cl/portal/url/item/bcb03d7bc28b64dfe040010165012d23.pdf (accessed on 11/May/2013).
20. Schmidt MTV. Chile: situación del tabaquismo a cinco años de la ratificación del Convenio Marco para el Control del Tabaco y los desafíos pendientes. http://www.chilelibredetabaco.cl/descargas/Informe_Chile_5_anos_CMCT_EPES_2010.pdf (accessed on 16/Jan/2013).
21. Medina E, Kaempffer AM. Tabaquismo y salud en Chile. Bol Oficina Sanit Panam 1991; 111:112-21.
22. Rojas G, Gaete J, González I, Ortega M, Figueroa A, Fritsch R, et al. Tabaquismo y salud mental. Rev Méd Chile 2003; 131:873-80.

Submitted on 12/Nov/2013
Final version resubmitted on 17/Dec/2014
Approved on 30/Mar/2015