



**Leticia Hachuel**

**Nora Arnesi**

**Gabriela Boggio**

**Marcos Prunello**

*Instituto de Investigaciones Teóricas y Aplicadas de la Escuela de Estadística*

## **COMPARACIÓN DE TENDENCIAS EN TASAS DE MORTALIDAD PARA DIFERENTES SUBPOBLACIONES MEDIANTE EL AJUSTE DE MODELOS EDAD-PERÍODO-COHORTE**

### **Introducción**

El análisis *edad-período-cohorte* se utiliza ampliamente en epidemiología para estudiar tendencias en las tasas de incidencia y de mortalidad por cáncer. Para llevar a cabo este análisis se cuenta con un conjunto de herramientas estándares que incluyen diferentes métodos que han ido evolucionando para atender a las necesidades crecientes de vigilancia epidemiológica del cáncer. Los métodos más utilizados incluyen gráficos basados en el diagrama de Lexis, el cálculo de tasas estandarizadas en forma directa por edad y la aplicación de métodos no paramétricos (Carstensen, 2007; Last, 1995; Tarone y Chu, 1996). Si bien estas herramientas estándares son especialmente adecuadas para análisis exploratorios y descriptivos resultan insuficientes para dar respuestas formales a muchos de los interrogantes planteados en el análisis *edad-período-cohorte*. Por este motivo, paralelamente en la literatura estadística se ha desarrollado el modelo estadístico *edad-período-cohorte* denominado modelo EPC (Fienberg y Mason, 1979; Clayton y Schifflers, 1987; Holford, 1983, 1991).

El modelo EPC se basa fundamentalmente en la teoría de los modelos lineales generalizados (McCullagh y Nelder, 1989; Agresti, 1986) y en principio, permite enriquecer el estudio descriptivo con la generación y prueba de hipótesis. Sin embargo, esta modelización es una metodología de uso poco frecuente en la práctica principalmente debido a la falta de una completa comprensión por parte del usuario del denominado problema de la identificación que afecta a los parámetros correspondientes a las variables intervinientes en el modelo.

Generalmente, un modelo EPC se ajusta a una matriz de tasas específicas según edad y período obtenida a partir de registros de cáncer basados en una población. Esta matriz de



tasas cuyas filas representan los distintos grupos de edad y las columnas los períodos bajo análisis, permite identificar las tasas específicas según cohorte utilizando la relación que establece que el año de nacimiento es igual al período calendario de diagnóstico o muerte menos la edad en el momento de diagnóstico o muerte. Esta relación  $cohorte = período - edad$  da origen al problema de la identificación antes mencionado.

Si bien no existe un consenso generalizado para la resolución de este problema, resulta indiscutible la necesidad de incluir las tres variables en la especificación del modelo. Esto se debe a que los efectos asociados a *edad* reflejan una historia natural subyacente relacionada con el envejecimiento, los efectos asociados a *período* capturan factores que influyen sobre las personas de todas las edades al mismo tiempo y los efectos asociados a cohorte persiguen el impacto neto de factores de riesgo cuya prevalencia varía de una generación a la siguiente.

En muchas situaciones el análisis EPC se lleva a cabo estratificando de acuerdo a subgrupos definidos por sexo, etnicidad, regiones geográficas, etc. Es posible, entonces, ajustar modelos EPC a diferentes conjuntos de tasas correspondientes a subpoblaciones y el objetivo del análisis se focaliza en la comparación y contraste de las respectivas tasas.

Bajo este enfoque, Rosenberg y Anderson (2010) presentaron pruebas de hipótesis lineales para determinar si las tasas definidas sobre las mismas edades, períodos y cohortes en dos subpoblaciones son absolutamente proporcionales o proporcionales luego de estratificar por *edad*, *período* o *cohorte*.

En el presente trabajo se evalúa si es posible hallar algún tipo de proporcionalidad en las tasas de mortalidad por cáncer de colon en Argentina entre hombres y mujeres para el período 1980-2009, siguiendo los lineamientos de trabajo de los autores recién citados.

## **Material y Métodos**

A partir de la información provista por los Registros de Estadísticas Vitales, organizada y publicada por el Ministerio de Salud de la Nación sobre la mortalidad en Argentina según causas de muerte (Codificación Diagnóstica Internacional, 10<sup>o</sup> revisión) y los Registros Censales (INDEC) es posible contar con información sobre el número de muertes por cáncer de colon para cada grupo de edad y período y sus respectivas cantidades de personas-año. Tradicionalmente dicha información se vuelca en tablas con tantas filas como grupos de edad considerados y tantas columnas como períodos bajo análisis.



Formalmente, sea  $\mathbf{Y} = \{y_{ep}, e = 1, \dots, E, p = 1, \dots, P\}$ , la matriz que contiene el número de muertes por cáncer en el período calendario  $p$  y el grupo etario  $e$  y  $\mathbf{N} = \{N_{ep}, e = 1, \dots, E, p = 1, \dots, P\}$ , la matriz que contiene el correspondiente número de personas-año.

Las diagonales de las matrices  $\mathbf{Y}$  y  $\mathbf{N}$ , consideradas a partir del extremo inferior izquierdo hasta el extremo superior derecho, representan las sucesivas cohortes de nacimiento indicadas por  $c = p - e + E$ , desde la cohorte más antigua ( $c = 1$ ) observada sólo para el mayor grupo de edad durante el primer período calendario ( $p = 1$ ), hasta la cohorte más joven ( $c = C = P + E - 1$ ) observada solamente para el intervalo de menor edad  $e = 1$  durante el período calendario más reciente  $p = P$ .

Las tasas observadas de mortalidad se obtienen a través del cociente  $\left\{ \frac{y_{ep}}{N_{ep}}; e = 1, \dots, E, p = 1, \dots, P \right\}$ , multiplicado por una constante elegida convenientemente para facilitar su interpretación, por ejemplo, 100000.

Un modelo lineal generalizado (MLG) adecuado para el estudio de las tasas de mortalidad se construye bajo los siguientes supuestos:

- El número de casos en el grupo de edad  $e$  y período de tiempo  $p$ ,  $y_{ep}$ , es una realización de una variable aleatoria Poisson con media  $\mu_{ep}$  ( $e = 1, \dots, E, p = 1, \dots, P$ ).
- El número de personas-año en el grupo de edad  $e$  y en el período  $p$ ,  $N_{ep}$ , es un valor fijo conocido.
- Las variables aleatorias  $y_{ep}$  son independientes.
- El logaritmo de la tasa esperada es función lineal de los efectos de los grupos de edad, períodos de tiempo y cohortes de nacimiento, es decir:

$$\ln \left( E \left( \frac{y_{ep}}{N_{ep}} \right) \right) = \ln \left( \frac{\mu_{ep}}{N_{ep}} \right) = \lambda + \alpha_e + \pi_p + \gamma_c \quad [1]$$

donde:

$\lambda$ : media general

$\alpha_e$ : efecto principal del  $e$ -ésimo grupo de edad;  $e = 1, \dots, E$ .



$\pi_p$ : efecto principal del  $p$ -ésimo período de tiempo;  $p = 1, \dots, P$ .

$\gamma_c$ : efecto principal de la  $c$ -ésima cohorte de nacimiento;  $c = 1, \dots, C = E + P - 1$ .

El subíndice  $c$  queda especificado al determinar un grupo de edad  $e$  y un período  $p$  a través de la relación  $c = p - e + E$ , motivo por el cual se utilizan sólo dos subíndices en  $y$ ,  $N$  y  $\mu$ .

En el ajuste de un MLG las restricciones que suelen emplearse son:

$$\sum_{e=1}^E \alpha_e = \sum_{p=1}^P \beta_p = \sum_{c=1}^C \gamma_c = 0,$$

o bien, considerar el efecto de un nivel de cada factor igual a cero, ya sea el primero, el último o cualquier otro que se desee tomar como referencia:

$$\alpha_e = \pi_p = \gamma_c = 0, \quad \text{para determinados } e, p \text{ y } c.$$

El logaritmo del número de personas-año en riesgo se asume como "offset" para el ajuste del modelo, ya que  $N_{ep}$  se obtiene de una estimación basada en datos censales cuya variación se considera no aleatoria y su coeficiente no debe ser estimado. De esta manera, el modelo [1] puede expresarse como:

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

El conjunto de restricciones habituales descrito anteriormente no soluciona el *problema de la identificación* que surge debido a la exacta relación lineal existente entre las variables: *cohorte = período - edad*. Es necesario, entonces, implementar una restricción adicional la cual podría tratarse, por ejemplo, de la igualdad de los efectos de dos niveles en uno de los factores.

Con las restricciones habituales y la restricción adicional impuesta es posible hallar un único conjunto de estimaciones máximo verosímiles de los parámetros del modelo. Sin embargo, los resultados que se obtienen dependen de esta restricción adicional, por lo que para su elección se requiere un profundo conocimiento del proceso estudiado, ya que si la misma no se corresponde con la realidad, se incurriría en interpretaciones sesgadas. A lo largo de los años se han desarrollado distintas soluciones al problema de la identificación de manera de evitar la arbitrariedad que en general conlleva dicha imposición. Entre las soluciones, una de las que goza de mayor popularidad es la propuesta por Holford (1983, 1991), que consiste



en limitar el análisis a combinaciones lineales o no lineales de los parámetros que permanecen invariantes con cualquiera de las restricciones adicionales que se imponga y que se las conoce bajo la denominación de "funciones estimables". Específicamente, este enfoque se centra en la partición de los efectos de las variables en dos componentes: una parte lineal y otra parte de curvatura o desviación de la linealidad. El modelo [1] puede expresarse de la siguiente forma:

$$\ln\left(\frac{\mu_{ep}}{N_{ep}}\right) = \lambda + \alpha_L \left(e - \frac{E+1}{2}\right) + \pi_L \left(p - \frac{P+1}{2}\right) + \gamma_L \left(c - \frac{C+1}{2}\right) + \tilde{\alpha}_e + \tilde{\pi}_p + \tilde{\gamma}_c \quad [2]$$

donde  $\alpha_L$ ,  $\pi_L$  y  $\gamma_L$  son las pendientes para la parte lineal,  $\tilde{\alpha}_e$ ,  $\tilde{\pi}_p$  y  $\tilde{\gamma}_c$  son las curvaturas o desviaciones de la linealidad y  $\left[e - \frac{E+1}{2}\right]$ ,  $\left[p - \frac{P+1}{2}\right]$  y  $\left[c - \frac{C+1}{2}\right]$  son los índices que resultan de centrar las variables *edad*, *período* y *cohorte*, respectivamente.

Las curvaturas constituyen funciones estimables, es decir satisfacen la condición de ser invariantes a la parametrización inicial. Por el contrario, las pendientes no lo son. A pesar de que las mismas pueden variar ampliamente con distintas restricciones adicionales, algunas combinaciones lineales de ellas son invariantes. Holford (1983) demuestra que siendo  $d_1$  y  $d_2$  constantes arbitrarias, la expresión:

$$d_1\alpha_L + d_2\pi_L + (d_2 - d_1)\gamma_L$$

es una función estimable de los parámetros.

Asignando valores a  $d_1$  y  $d_2$ , se tienen las siguientes funciones estimables, cada una de las cuales recibe un nombre particular:

- $d_1 = d_2 = 1 \Rightarrow \alpha_L + \pi_L$  es estimable  $\Rightarrow$  *tendencia longitudinal de edad*
- $d_1 = 1, d_2 = 0 \Rightarrow \alpha_L - \gamma_L$  es estimable  $\Rightarrow$  *tendencia transversal de edad*
- $d_1 = 0, d_2 = 1 \Rightarrow \pi_L + \gamma_L$  es estimable  $\Rightarrow$  *drift neto*

Se puede observar que la *tendencia transversal de edad* es la diferencia entre la *tendencia longitudinal de edad* y el *drift neto*:

$$(\alpha_L + \pi_L) - (\pi_L + \gamma_L) = \alpha_L - \gamma_L.$$

Los resultados anteriores son de interés porque sugieren que, si bien las pendientes no son estimables y no están delimitadas, ellas no pueden variar independientemente una de otra.



Si una cualquiera de ellas se fija en algún valor particular, entonces las otras dos quedan también determinadas.

De acuerdo a estas definiciones el modelo [2] puede reescribirse en términos de las funciones estimables de dos formas diferentes, conocidas como edad-cohorte [3] y edad-período [4]:

$$\ln\left(\frac{\mu_{ep}}{N_{ep}}\right) = \lambda + (\alpha_L + \pi_L)\left(e - \frac{E+1}{2}\right) + (\pi_L + \gamma_L)\left(c - \frac{C+1}{2}\right) + \tilde{\alpha}_e + \tilde{\pi}_p + \tilde{\gamma}_c \quad [3]$$

$$\ln\left(\frac{\mu_{ep}}{N_{ep}}\right) = \lambda + (\alpha_L - \gamma_L)\left(e - \frac{E+1}{2}\right) + (\pi_L + \gamma_L)\left(p - \frac{P+1}{2}\right) + \tilde{\alpha}_e + \tilde{\pi}_p + \tilde{\gamma}_c \quad [4]$$

Si el objetivo del estudio es comparar las tasas de mortalidad en dos subpoblaciones, por ejemplo hombres y mujeres, se tienen dos conjuntos independientes de tasas de mortalidad, una para cada subpoblación, definidas sobre los mismos grupos de edad  $e = 1, \dots, E$ , períodos  $p = 1, \dots, P$  y cohortes  $c = 1, \dots, P + A - 1$ .

Bajo el supuesto que un modelo EPC ajusta adecuadamente en cada subpoblación, la metodología propuesta por Rosenberg y Anderson (2010) tiene como objetivo establecer la relación entre ambos modelos buscando particularmente evidencia de algún tipo de proporcionalidad en las tasas, la cual puede manifestarse como una proporcionalidad absoluta – para todo grupo de edad y período– o proporcionalidad luego de estratificar por edad, período o cohorte. Los autores desarrollaron cuatro tipos de proporcionalidad, definidos según qué parámetros de los respectivos modelos EPC resulten iguales para las dos subpoblaciones.

Si se supra indica con  $j$  ( $j = 0, 1$ ) al modelo EPC según sea para una subpoblación u otra, los autores definen la *proporcionalidad absoluta* entre las tasas de ambos grupos si resulta:

$$\ln\left(\frac{\mu_{ep}^1}{N_{ep}^1}\right) = \phi + \ln\left(\frac{\mu_{ep}^0}{N_{ep}^0}\right)$$

La cantidad  $\exp(\phi)$  es por lo tanto la razón de tasas, también denominada razón de riesgos, para el grupo 1 respecto al grupo 0, y es constante a través de las edades, períodos y cohortes. Por lo tanto, para que se verifique la proporcionalidad definida deben cumplirse las siguientes igualdades entre los parámetros:



$$\begin{aligned}
 (\alpha_L + \pi_L)^1 &= (\alpha_L + \pi_L)^0 \\
 (\pi_L + \gamma_L)^1 &= (\pi_L + \gamma_L)^0 \\
 \tilde{\alpha}_e^1 &= \tilde{\alpha}_e^0 & e = 1, \dots, E \\
 \tilde{\pi}_p^1 &= \tilde{\pi}_p^0 & p = 1, \dots, P \\
 \tilde{\gamma}_c^1 &= \tilde{\gamma}_c^0 & c = 1, \dots, C
 \end{aligned}$$

de donde se desprende que  $\phi = \lambda^1 - \lambda^0$ , no dependiendo de  $p$ ,  $e$  ni  $c$ .

De manera similar se verifica la denominada *proporcionalidad dentro de cohorte* cuando la tendencia longitudinal de edad, las desviaciones de edad y las desviaciones de período son iguales en ambos grupos. Esto es:

$$\begin{aligned}
 (\alpha_L + \pi_L)^1 &= (\alpha_L + \pi_L)^0 \\
 \tilde{\alpha}_e^1 &= \tilde{\alpha}_e^0 & e = 1, \dots, E \\
 \tilde{\pi}_p^1 &= \tilde{\pi}_p^0 & p = 1, \dots, P
 \end{aligned}$$

Existe *proporcionalidad dentro de grupos de edad* si los logaritmos de las tasas de ambos grupos difieren en una constante que sólo depende del grupo de edad en consideración, igual a través de todos los períodos. Es decir, deben verificarse las siguientes igualdades en los parámetros:

$$\begin{aligned}
 (\pi_L + \gamma_L)^1 &= (\pi_L + \gamma_L)^0 \\
 \tilde{\pi}_p^1 &= \tilde{\pi}_p^0 & p = 1, \dots, P \\
 \tilde{\gamma}_c^1 &= \tilde{\gamma}_c^0 & c = 1, \dots, C
 \end{aligned}$$

En otras palabras, este tipo de proporcionalidad implica que son iguales en ambos grupos el drift neto, las desviaciones de período y las desviaciones de cohorte.

Por último, puede presentarse la *proporcionalidad dentro de períodos* si los logaritmos de las tasas de ambos grupos difieren en una constante que sólo depende del período en consideración, igual a través de todas las edades. Las igualdades en los parámetros que deben verificarse son:

$$\begin{aligned}
 (\alpha_L - \gamma_L)^1 &= (\alpha_L - \gamma_L)^0 \\
 \tilde{\alpha}_e^1 &= \tilde{\alpha}_e^0 & e = 1, \dots, E \\
 \tilde{\gamma}_c^1 &= \tilde{\gamma}_c^0 & c = 1, \dots, C
 \end{aligned}$$

Luego, este tipo de proporcionalidad implica la igualdad en ambos grupos de la tendencia



transversal de edad, las desviaciones de edad y las desviaciones de cohorte.

Para la construcción de los tests que evalúan estos tipos de proporcionalidad es conveniente utilizar los vectores de parámetros del modelo EPC explicitado en la forma *edad-cohorte* (modelo [3]) y en la forma *edad-período* (modelo [4]). Esto es:

$$\text{Forma edad-cohorte: } \boldsymbol{\psi}^j = \left( \lambda^j, (\alpha_L + \pi_L)^j, (\pi_L + \gamma_L)^j, \tilde{\boldsymbol{\alpha}}_{(2:E-1)}^j, \tilde{\boldsymbol{\pi}}_{(2:P-1)}^j, \tilde{\boldsymbol{\gamma}}_{(2:C-1)}^j \right)$$

$$\text{Forma edad-período: } \boldsymbol{\psi}^j = \left( \lambda^j, (\alpha_L - \gamma_L)^j, (\pi_L + \gamma_L)^j, \tilde{\boldsymbol{\alpha}}_{(2:E-1)}^j, \tilde{\boldsymbol{\pi}}_{(2:P-1)}^j, \tilde{\boldsymbol{\gamma}}_{(2:C-1)}^j \right)$$

Los estimadores de los parámetros tienen distribución asintótica normal y es posible obtener su matriz de variancias y covariancias  $\text{Var}(\hat{\boldsymbol{\psi}}^j)$  de manera sencilla. Cada uno de las diferentes proporcionalidades resultan consistentes con los datos si se verifican las igualdades antes definidas entre parámetros de los vectores  $\boldsymbol{\psi}^0$  y  $\boldsymbol{\psi}^1$ .

Bajo el supuesto de independencia entre las dos subpoblaciones, se pueden construir tests tipo Wald para comprobar cada una de las hipótesis de proporcionalidad, que toma la forma general:

$$X^2 = \left[ \mathbf{L}' (\hat{\boldsymbol{\psi}}^1 - \hat{\boldsymbol{\psi}}^0) \right]' \left[ \mathbf{L}' (\text{Var}(\hat{\boldsymbol{\psi}}^1) - \text{Var}(\hat{\boldsymbol{\psi}}^0)) \mathbf{L} \right]^{-1} \left[ \mathbf{L}' (\hat{\boldsymbol{\psi}}^1 - \hat{\boldsymbol{\psi}}^0) \right]$$

donde  $\mathbf{L}$  es una matriz de unos y ceros según los parámetros que se necesitan comparar en las diferentes hipótesis (Searle, 1971).

Para la comprobación de dichas hipótesis de proporcionalidad se contrasta el valor  $X^2$  con la distribución  $\chi^2$  con la cantidad de parámetros considerados iguales en cada hipótesis como grados de libertad.

## Resultados

Se construyeron dos conjuntos de tasas de mortalidad por cáncer de colon en la Argentina, para hombres y mujeres, considerando grupos quinquenales de edad definidos entre 25 y 79 años y para el período 1980-2009 agrupado en quinquenios.

En una primera instancia, se analiza a través de un examen gráfico la posible existencia de proporcionalidad en las tasas de mortalidad por cáncer de colon entre las dos subpoblaciones: hombres y mujeres.

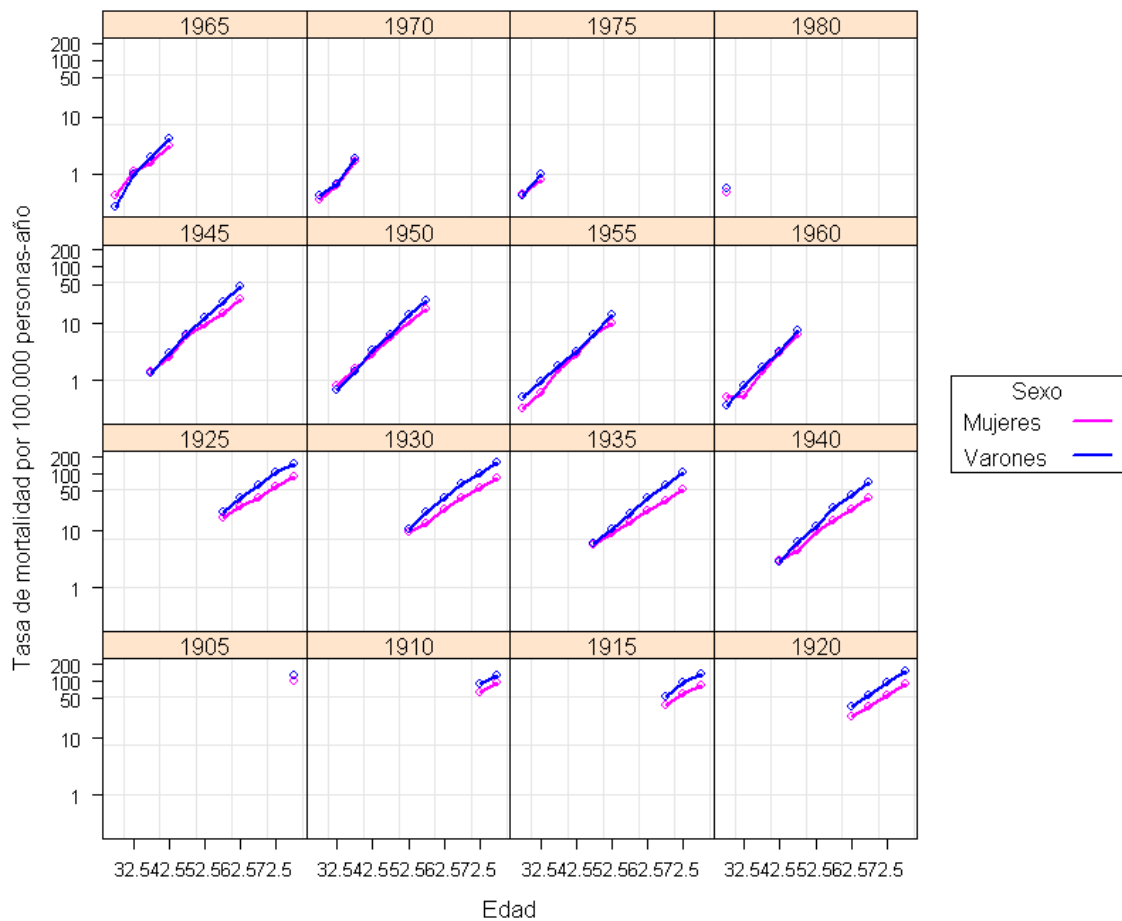




La Figura 1 compara las tasas de mortalidad según la cohorte de nacimiento entre ambas subpoblaciones. Cada cohorte fue representada en un panel diferente para evitar que los solapamientos impidan la fácil visualización. Este tipo de gráfico sirve para detectar proporcionalidad dentro de cohortes; en tal caso las curvas para hombres y mujeres resultan aproximadamente paralelas aunque el espacio entre las mismas puede cambiar dependiendo de la cohorte que se considere. Si la separación para todas las cohortes fuese la misma se tendría un indicio de proporcionalidad absoluta.

Concretamente, en este estudio parecería que no hay indicios de proporcionalidad dentro cohortes puesto que, en general, la brecha entre las curvas es menor en edades más jóvenes y se amplía con el envejecimiento de la cohorte. Por esta misma razón, se debe descartar también la existencia de proporcionalidad absoluta.

**Figura 1. Tasas de Mortalidad versus Edad, según Cohorte**

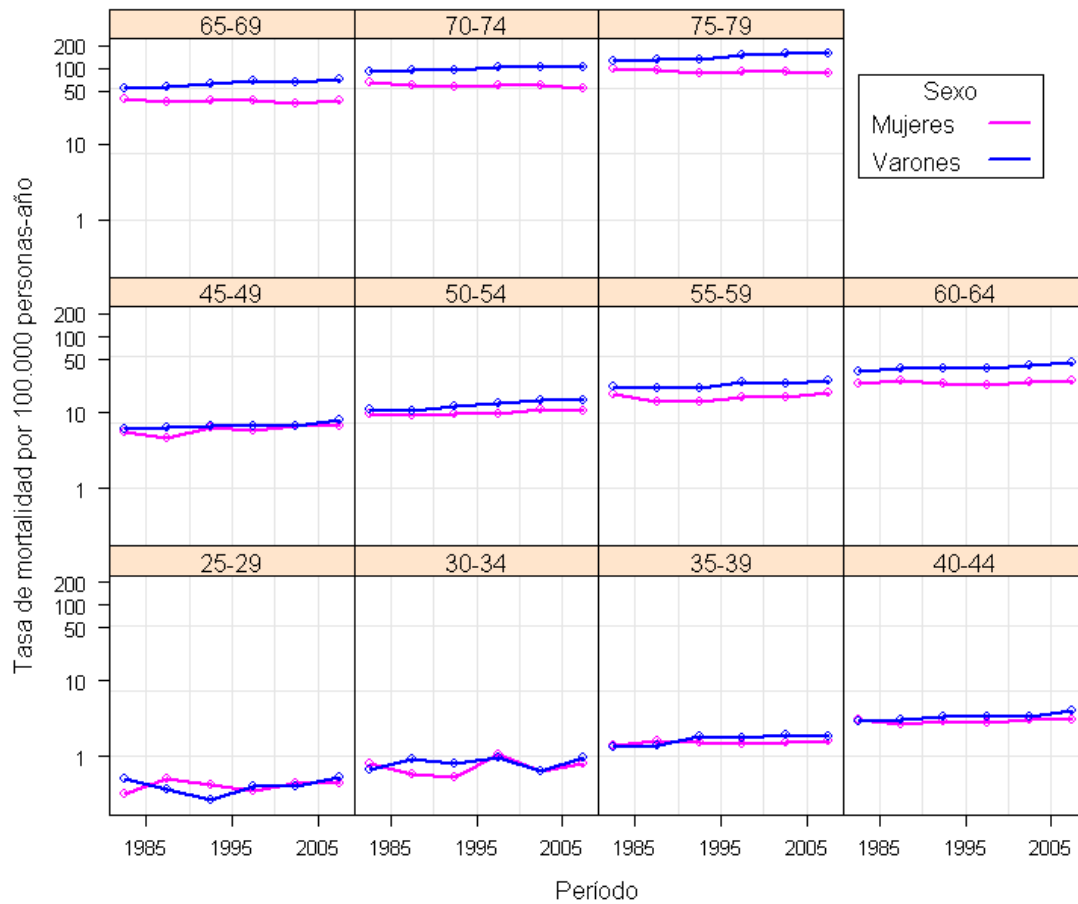


La Figura 2 presenta las tasas específicas por edad a través de los períodos estudiados. Es



posible observar cierta evidencia de proporcionalidad en las tasas entre grupos de edad, principalmente para las edades superiores ya que las curvas evidencian cierto paralelismo entre ambos grupos.

**Figura 2. Tasas de Mortalidad versus Período, según Edad**

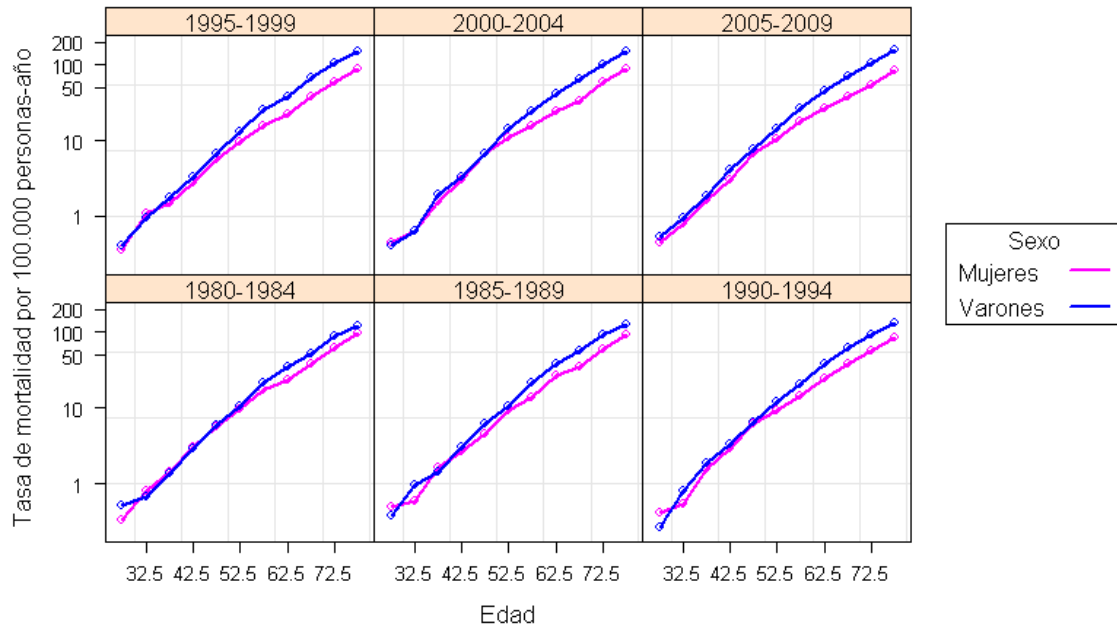


Por último, la Figura 3 compara las tasas de mortalidad entre ambos sexos según el período calendario. Este tipo de gráfico sirve para detectar proporcionalidad dentro de los períodos, si se observa que las curvas de ambos sexos son aproximadamente paralelas y el espacio entre las mismas puede cambiar en función del período que se considere. Si la separación entre las curvas de los varones y mujeres fuese la misma para todos los períodos, se tendría un indicio de proporcionalidad absoluta.

En este caso particular no se detectan signos de proporcionalidad dentro de períodos puesto que, en general, la brecha entre las curvas de cada sexo es menor en edades jóvenes y se amplía en edades superiores.



**Figura 3. Tasas de Mortalidad versus Edad, según Período**



En una segunda etapa se ajustaron modelos EPC en cada subpoblación; en el correspondiente a las mujeres se detectó sobredispersión, por lo cual se aplicó la corrección apropiada logrando de esta manera un ajuste adecuado tanto para las tasas de mortalidad masculinas como femeninas.

La Tabla 1 muestra los resultados de los tests de Wald construidos para las cuatro hipótesis de proporcionalidad.

**Tabla 1. Resultados de los Tests de proporcionalidad**

Hipótesis	GL	Estadística de Wald	Prob. asociada
Prop. Absoluta	29	443.87	< 0.0001
Prop. dentro de cohortes	14	219.34	< 0.0001
Prop. dentro de grupos de edad	19	163.09	< 0.0001
Prop. dentro de períodos	24	283.32	< 0.0001



Todos los tests resultaron significativos, por lo cual se concluye que para el cáncer de colon en Argentina, no se evidencia proporcionalidad de ninguno de los tipos estudiados en las tasas de mortalidad masculinas y femeninas. Ello, dicho de otro modo, se interpreta como que la razón de tasas entre hombres y mujeres es heterogénea cualquiera sea la estratificación considerada y con menor énfasis dentro de grupos de edad (en concordancia con el valor más chico de la estadística de Wald). Este resultado no sorprende dado que es común en la práctica no encontrar proporcionalidad entre las tasas referidas a subpoblaciones relacionadas con el género.

## Discusión

Los estudios comparativos relacionados con tasas de mortalidad por cáncer se basan generalmente en estudios de carácter descriptivo. Este enfoque resulta conveniente para una fácil visualización de las tendencias en las tasas de mortalidad entre subpoblaciones. Sin embargo presenta complicaciones por el hecho de considerar tres escalas: *edad*, *período* y *cohorte* interrelacionadas y considerar sólo dos de ellas en los gráficos. Por lo tanto, a priori no resulta claro cuál de las representaciones gráficas revela más claramente las tendencias indagadas.

En este trabajo se aplica una propuesta metodológica de características más formales presentada por Rosenberg y Anderson (2010), útil a la hora de organizar el estudio comparativo. Concretamente es posible evaluar distintos tipos de proporcionalidad estableciendo condiciones de igualdad entre ciertos parámetros específicos entre los modelos EPC ajustados a las subpoblaciones. El resultado de esta comparación de parámetros indica cuál es el gráfico más apropiado para la descripción el comportamiento de las tasas.

Aún cuando ningún tipo de proporcionalidad se compruebe, como es el caso presentado en este trabajo, las tasas estimadas a través de los modelos EPC siempre resultan versiones suavizadas de las tasas crudas permitiendo diferenciar los aspectos más importantes de la perturbación.

Específicamente, en el estudio sobre el cáncer de colon en Argentina, la utilización del método formal permitió concluir acerca de la heterogeneidad de las razones de tasas de mortalidad entre hombres y mujeres estratificando por grupo de *edad*, *período* o *cohorte*.



## REFERENCIAS BIBLIOGRÁFICAS

- Agresti A. (1986). *An Introduction to Categorical Data Analysis*. Wiley Series in Probability and Statistics. New York: Wiley.
- Carstensen B. (2007). Age-period-cohort models for Lexis diagram. *Statistics in Medicine*; 26:3018-45.
- Clayton D., Schifflers E. (1987). Models for temporal variation in cancer rates II: age-period-cohort models. *Statistics in Medicine*; 6: 469-481.
- Fienberg S.E., Mason W. N. (1979). Identification and estimation of age-period-cohort in the analysis of discrete archival data". *Sociological Methodology*; 1-67. Schuessler K.F. San Francisco. Jossey-Bass.
- Holford T.R. (1983). The estimation of age, period and cohort effects for vital rates. *Biometrics*; 39:311-324.
- Holford T.R. (1991). Understanding the effects of Age, Period, and Cohort on incidence and mortality rates. *Annual Review of Public Health*; 12: 425-457.
- Holford T.R. (1992). Analysing the temporal effects of age, period and cohort. *Statistical Methods in Medical Research*; 1:317-337.
- McCullagh P., Nelder J.A. (1989). *Generalized linear models*. 2<sup>nd</sup> ed. New York: Chapman & Hall; 1989.
- Rosenberg P.S., Anderson, W.F. (2010). Proportional hazards models and age-period-cohort analysis of cancer rates. *Statistics in Medicine*; 29:1228-1238
- Searle S.R. (1971). *Linear Models*. New York: John Wiley & Sons, Inc.
- Tarone RE, Chu K.C. (1996). Evaluation of birth cohort patterns in population disease rates. *American Journal of Epidemiology* ;143:85-91.

## FUENTES

- Anuarios. Ministerio de Salud de la República Argentina. [www.msal.gov.ar](http://www.msal.gov.ar). Acceso el 10 de marzo de 2010.
- Estimaciones y Proyecciones de Población. Total del país. 1950-2015. Instituto Nacional de Estadísticas y Censos (INDEC) [www.indec.mecon.gov.ar](http://www.indec.mecon.gov.ar). Acceso el 10 marzo de 2010.