

# Métodos de análisis de estudios prospectivos laborales: la estimación de la razón de mortalidad estandarizada

Fernández, F. y Kogevinas, M.

## RESUMEN

El presente artículo es una revisión de la metodología estadística para el análisis de estudios de cohortes laborales y pretende ofrecer información sobre posibles fuentes de registros de mortalidad y algunos programas y paquetes estadísticos disponibles para realizar este tipo de análisis. Hacemos énfasis en los métodos para realizar una evaluación cuantitativa de la magnitud del riesgo asociado a la exposición. Nos referimos con más detalle a uno de los métodos más frecuentemente aplicados, el cálculo de la razón de mortalidad estandarizada (RME). Otros conceptos tratados son la selección de grupos de comparación poblacionales externos e internos, el concepto «personas-tiempo a riesgo», el periodo de latencia para la aparición de la enfermedad, y los métodos de evaluación indirecta de factores de confusión (p. ej., el tabaco). Como ejemplo, se presentan todos los pasos de un análisis de una pequeña cohorte de plomistas españoles expuestos al amianto.

## PALABRAS CLAVE

Razón de mortalidad estandarizada. Epidemiología laboral. Estudios de cohorte. Métodos estadísticos. Factores de confusión.

## STATISTICAL METHODS FOR THE ANALYSIS OF PROSPECTIVE OCCUPATIONAL STUDIES: ESTIMATION OF THE STANDARDIZED MORTALITY RATIO

### SUMMARY

This paper reviews the statistical methods for data analysis of occupational cohort studies and gives information on sources of mortality records, and information on some programs and statistical packages to be used in this type of analyses. We focus on methods to derive quantitative risk estimation associated with a specific exposure. We refer in detail to one of the most frequently used methods, the standardised mortality ratio (SMR). Selection of external and internal comparison groups, the term «persons-time at risk», the use of latency periods for the occurrence of disease and indirect methods to assess the confounding effect of tobacco are also discussed. We illustrate all the steps in the data analysis using as an example a small cohort of Spanish pipefitters exposed to asbestos.

### KEY WORDS

Standardized mortality ratio. Occupational epidemiology. Cohort studies. Statistical methods. Confounding factors.

El diseño de estudio de cohortes es uno de los más frecuentemente aplicados en epidemiología laboral, especialmente en estudios que investigan enfermedades crónicas. Este tipo de estudios se han aplicado para determinar, por ejemplo, la cancerogenicidad de la b-naftilamina, el benceno, el amianto, las dioxinas o la presencia de asma laboral en panaderos o agricultores<sup>1-3</sup>. En España se han estudiado cohortes en trabajadores de la industria del papel y trabajadores de centrales nucleares<sup>4,5</sup>, en los cuales se examinan poblaciones de sectores industriales específicos.

En este trabajo revisaremos los elementos principales de los estudios de cohortes y utilizaremos el análisis de una cohorte de plomistas de una compañía de gas para ilustrar dichos principios. Esta segunda parte del trabajo se presenta en los cuadros que siguen.

### Diseño del estudio

En los estudios de cohorte, una vez se ha identificado un grupo de individuos empleados en un sector, se recoge la información sobre su exposición laboral y se siguen a lo largo del tiempo para conocer la frecuencia de la enfermedad de interés. Para cada individuo, la información sobre su exposición laboral previa se relaciona con la subsiguiente ocurrencia de enfermedad. La tasa de incidencia o mortalidad de la enfermedad en el grupo de trabajadores expuestos se compara con la de otro grupo de no expuestos.

Normalmente, se hace la distinción entre estudios prospectivos de cohorte y estudios retrospectivos de cohorte. En los estudios prospectivos de cohorte se recoge

Unitat de Recerca Respiratòria i Ambiental. Institut Municipal d'Investigació Mèdica. Barcelona.

Correspondencia: F. Fernández. Unitat de Recerca Respiratòria i Ambiental. Institut Municipal d'Investigació Mèdica. Doctor aiguader, 80. 08003 Barcelona. Correo electrónico: Ffernandez@imim.es

Trabajo recibido el 4-1-01. Aceptado el 7-5-01.

información sobre la exposición individual y los sujetos son seguidos prospectivamente en el tiempo. Los estudios históricos de cohortes (estudios retrospectivos de cohortes) sólo se pueden realizar si existen registros completos de los trabajadores de una industria y se dispone de registros completos de mortalidad (o morbilidad) de la población general. En los estudios retrospectivos de cohorte se identifican poblaciones expuestas y no expuestas a un factor laboral en un momento definido del pasado y se determina el número de individuos que ha desarrollado la enfermedad en estudio.

En nuestro ejemplo, la población de estudio (cohorte laboral) es dinámica; es decir, hay trabajadores que entran en el estudio en un tiempo posterior al inicio de éste. En otras ocasiones, cuando una cohorte está establecida después de una accidente industrial, ésta puede ser fija, es decir, todos los miembros de la cohorte entran en el período de seguimiento del estudio en la misma fecha. También existen trabajadores que abandonan o son «perdidos» durante el período de estudio. Según estos acontecimientos, la duración del tiempo de seguimiento de cada sujeto es diferente, tal y como se refleja en la figura 1.

Cuadro 1. Descripción de la cohorte de plomistas de una compañía de gas

La cohorte incluye 44 varones que han trabajado en una industria del gas en Barcelona durante un periodo no inferior a un año como plomistas, instalando y reparando tuberías en áreas públicas. Los primeros plomistas comenzaron a trabajar en la compañía en enero de 1964 y el periodo de estudio empieza en esta fecha. El último plomista empezó a trabajar y entró en el estudio el 15 de enero de 1998. Los plomistas han sido seguidos para evaluar su mortalidad hasta el 31 de diciembre de 1999 o hasta la fecha de fallecimiento. La media de edad al inicio del empleo fue de 30,7 años, con una desviación estándar (DE) de 7,2. La media de edad al finalizar el período de seguimiento fue de 52,4 años, con una DE de 13,0. A su vez, la media de duración del empleo fue de 11,9 años con una DE de 9,7. Se contactó con los 44 trabajadores o con sus respectivas familias, personalmente o por teléfono, para certificar su estado vital. Durante el periodo de seguimiento fallecieron 4 trabajadores. Inicialmente, la causa de muerte fue identificada tras contactar con la familia y, posteriormente, verificada mediante la revisión de registros hospitalarios

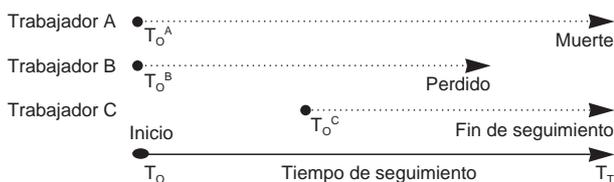


Figura 1. Ejemplos del tiempo de seguimiento en trabajadores de una cohorte laboral.

Tasas de incidencia/mortalidad. El concepto de personas-años

El método básico utilizado para estimar tasas de mortalidad en una cohorte se basa en el cálculo de lo que se ha denominado «personas-tiempo a riesgo», que se obtiene sumando los tiempos de seguimiento contribuidos por cada individuo que permanece bajo observación y que está a riesgo de convertirse en un «caso». Por tanto, cada individuo contribuye con unas personas-años de observación desde el inicio del seguimiento hasta que éste desarrolla el episodio de interés, es perdido o el tiempo de seguimiento finaliza.

Cuadro 2. Ejemplo de variables temporales de un individuo en la cohorte estudiada

	Fecha	Edad individuo	
Entrada	15/11/1964	(26 años)	Las personas-años aportadas por este individuo fueron 35,123 (la diferencia entre 31/12/1999 y 15/11/1964)
Salida	15/10/1996	(58 años)	
Fin de seguimiento	31/12/1999	(61 años)	

Generalmente, las personas-tiempo contribuidas colectivamente por toda la cohorte puede estratificarse de acuerdo con varios factores vinculados al tiempo<sup>6</sup>, tales como la edad de riesgo, el año o periodo, el tiempo de duración de ocupación, el tiempo desde el inicio de la exposición, la duración del seguimiento, la edad de la primera exposición y el año de la primera exposición, tal y como se refleja en la figura 2.

Sin embargo, de todos estos factores, la edad y el periodo son las variables de estratificación más comunes en los estudios de cohortes, puesto que el riesgo de muchas enfermedades varía con la edad y, además, existen enfermedades con tendencias definidas, así como exposiciones de diferente intensidad, relacionadas con el año. Por otro lado, este tipo de estratificación facilita la comparación de las tasas de mortalidad de registros nacionales con poblaciones externas, eliminando así posibles efectos confundidores de estos dos factores.

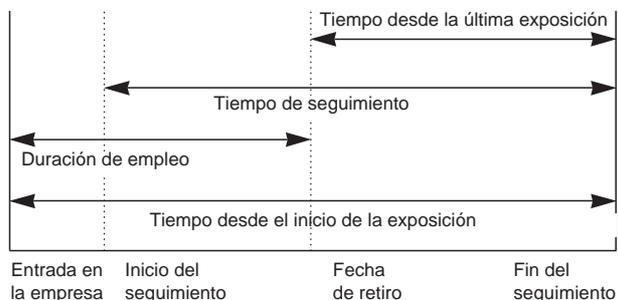


Figura 2. Factores asociados con el tiempo y su interrelación, analizados en estudios de cohortes laborales.

Una herramienta gráfica muy útil para representar de manera individual el concepto de personas-tiempo a riesgo es el diagrama de Lexis, representado en la figura 3. Se trata de un diagrama bidimensional compuesto por celdas que representan estratos de edad y periodos de tiempo, y que nos permiten estimar la contribución de cada individuo, desde su entrada en la cohorte hasta el fin de período de seguimiento<sup>7</sup>. Cada individuo se desplaza diagonalmente según sus cambios de edad y período, contribuyendo así al cómputo de personas-tiempo a riesgo por cada uno de los estratos a través de los cuales se desplaza. Las tasas específicas de mortalidad para cada estrato se calculan dividiendo el número de casos declarados en cada estrato de edad-período por el factor personas-tiempo a riesgo correspondiente, tal y como se expresa en la tabla 1.

En el cuadro 3 se representa la contribución de las personas-años en cada uno de los estratos de edad y período de tiempo de la persona mencionada en el cuadro 2.

Cuadro 3. Ejemplo del cálculo de la contribución de personas-años de un individuo en la cohorte estudiada en cada uno de los estratos de edad y período de tiempo

Período de tiempo	Grupo de edad	Personas-años
1964-1968	20-34	4,123
1969-1973	20-34	4,58
	35-44	0,42
1974-1978	35-44	5
1979-1983	35-44	4,58
	45-54	0,42
1984-1988	45-54	5
1989-1993	45-54	4,58
	55-64	0,42
1994-1998	55-64	5
1999-2000	55-64	1
Total		35,123

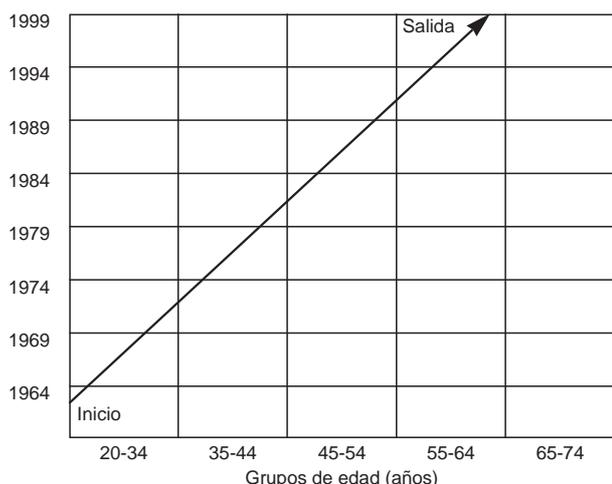


Figura 3. Diagrama de Lexis, donde se representan las personas-años de un sujeto de la cohorte, distribuidas por edad y período de tiempo.

Tabla 1. Fórmulas para el cálculo de las tasas específicas de mortalidad y de la razón de mortalidad estandarizada (RME)

<p>Tasa específica de mortalidad (por 100.000) =</p> $\frac{\text{N.º de casos (muertes) en la población definida, en un grupo de edad y período de tiempo específicos}}{\text{Personas-años a riesgo en esta población, en este grupo de edad y durante este mismo período de tiempo}} \times 100.000$
<p>Razón de mortalidad estandarizada (RME) =</p> $\frac{\text{N.º de casos (muertes) observados en una población definida}}{\text{N.º de casos (muertes) esperadas* en esta población, si siguiera la misma distribución de tasas específicas de mortalidad que una población «estándar»}}$
<p>Donde, *N.º de muertes esperadas =</p> $\sum \left( \text{Tasa específica en cada grupo de edad de la población «estándar»} \times \text{personas-años a riesgo en el correspondiente grupo de edad de nuestra población} \right)$

Siguiendo los mismos principios, se calculan las personas-años de toda la cohorte. Este cálculo se puede hacer utilizando diferentes programas estadísticos desarrollados especialmente para este tipo de análisis, por ejemplo: PAMCOMP 1.1<sup>8</sup>, OCMAP-PLUS 2.0<sup>9</sup>, PC-LTAS 1.0c<sup>10</sup>, o escribiendo algoritmos específicos para su utilización con paquetes estadísticos de más amplia utilización, como STATA 6.0<sup>11</sup>. Para obtener más información sobre el programa utilizado para el cálculo de las personas-años y la estimación de la RME (IC del 95%) utilizando STATA 6.0, se puede contactar con el autor.

(v. cuadro 4)

### Selección de grupos de comparación

En los estudios de cohortes se utilizan fundamentalmente dos tipos de grupos de comparación: a) grupos de comparación externos (frecuentemente tasas nacionales de mortalidad, o en ocasiones la mortalidad de otra cohorte de trabajadores, y b) grupos de comparación internos, que se identifican dentro de la misma cohorte. Ocasionalmente, se utilizan ambos tipos de grupos de comparación en un mismo estudio.

Cuando se analizan las tasas de mortalidad de forma global, la práctica más común es la de escoger la población nacional como grupo de comparación y calcular la razón de mortalidad estandarizada (RME). Existen diversas ventajas al utilizar la población nacional como población de referencia. En primer lugar, tiene interés conocer cómo se comportan las tasas de la cohorte res-

Cuadro 4. Estimación de personas año en la cohorte de plomistas

Para cada trabajador se ha calculado las personas-años a partir de un año después del inicio del empleo, hasta la ocurrencia del episodio (muerte) o fin de seguimiento (31 de diciembre de 1999). Un total de 910,11 personas-años a riesgo fueron acumuladas durante el período de seguimiento (1964-1999). La razón por la cual se han calculado las personas-años a partir de un año después del empleo es porque el criterio para formar parte de la cohorte era el empleo durante al menos un año como plomista en esta empresa. Por definición, para ser miembro de la cohorte una persona tenía que «sobrevivir» este primer año, que quiere decir que (para el estudio) no estaba en riesgo de morir. Si hubiera muerto durante el primer año no habría podido ser miembro de la cohorte. Las personas que trabajaron menos tiempo estaban excluidas

Cómputo global de personas-años a riesgo en la cohorte de estudio durante 1964-1999, por edad y período de tiempo

Estrato de edad (años)	Períodos de tiempo									
	1963-1964	1965-1966	1967-1968	1969-1970	1971-1972	1973-1974	1975-1976	1977-1978	1979-1980	1981-1982
20-34	0	12,58	14,98	13,62	13,01	9,13	13,54	13,72	8,96	5,48
35-44	0	3,13	7,38	14,06	12,1	19,17	23,48	22,88	23,23	20,89
45-54	0	0	0	0,86	5,45	6,87	9,17	13,36	16,98	18,19
55-64	0	0	0	0	0	0	0	0	0,87	5,45
65-74	0	0	0	0	0	0	0	0	0	0
(años)	1983-1984	1985-1986	1987-1988	1989-1990	1991-1992	1993-1994	1995-1996	1997-1998	1999-2000	
20-34	4,55	8,31	12,39	12,97	14,54	14,8	13,03	14,38	5,98	
35-44	15,72	13,59	17,36	15,39	15,19	15,11	18,43	12,57	7,1	
45-54	25,43	24,51	21,93	21,19	19,88	13,69	13,56	19,3	10,36	
55-64	6,88	9,16	14,37	18,97	18,6	22,08	21	16,7	7,66	
65-74	0	0	0	0,86	5,45	6,87	9,17	13,71	8,83	
										Total
										910,11

pecto a las tasas nacionales. En segundo lugar, las tasas de mortalidad por grupos de edad, sexo y período son habitualmente accesibles tanto a nivel nacional a través del Instituto Nacional de Estadística<sup>12</sup>, ya sea mediante publicaciones o por soporte informático, como a nivel internacional<sup>13</sup>. Estos datos de mortalidad están codificados según la Clasificación Internacional de Enfermedades proporcionada por la Organización Mundial de la Salud, y que son revisadas cada 10 años<sup>14</sup>. Recientemente, se ha establecido el Índice Nacional de Defunciones (IND), que pretende ser un punto de referencia a nivel nacional, proporcionando datos sobre el estado vital, aunque carece de información referente a las causas de muerte.

En términos de incidencia de cáncer a nivel internacional, una importante fuente de información sobre registros poblacionales se encuentra en el volumen «Cancer Incidence in five Continents»<sup>15</sup>.

Por último, el uso de tasas de mortalidad nacionales tiene la ventaja estadística de que son numéricamente estables, lo cual es importante cuando se investigan enfermedades poco frecuentes. Sin embargo, la utilización de las tasas de mortalidad nacionales tiene algunas desventajas. En estudios que investigan poblaciones laborales las comparaciones pueden estar sesgadas

debido al «efecto del trabajador sano». Éste se refiere a que la población activa es más sana que la población inactiva por el simple hecho de que personas con problemas de salud se excluyen del mercado del trabajo. Sin embargo, en el cálculo de las tasas nacionales se incluyen tanto las tasas de la población activa como las de parte de la población inactiva. Por esta razón, en muchas cohortes laborales las tasas de mortalidad global son más bajas que las correspondientes a la población nacional. Otra desventaja es la dificultad de controlar por factores potencialmente confusores que no han sido medidos, como los relacionados con el estilo de vida o las exposiciones ambientales. Estos problemas pueden ser parcialmente eliminados si se utilizan, como población de referencia, otras poblaciones externas definidas. En este caso se pueden realizar comparaciones con las tasas de la población activa, como ocurre en los países escandinavos, con poblaciones regionales o locales, o con otras cohortes externas (p. ej., otras poblaciones industriales). El uso de las tasas regionales está indicado cuando se sabe que la morbilidad o la mortalidad varían considerablemente entre distintas regiones. Por ejemplo, en Cataluña, la mortalidad por cáncer de pulmón es aproximadamente un 20% mayor que en el resto de España<sup>16</sup>. Ello se puede

deber a la variabilidad de la prevalencia de exposiciones laborales, pero también a diferencias en estilos de vida o en exposiciones ambientales. Por tanto, un estudio que investigue la mortalidad por cáncer de pulmón en una industria localizada en Cataluña tenderá a sobrestimar el riesgo de morir si se utilizan como referencia las tasas nacionales españolas. Las principales ventajas de utilizar como población de referencia una cohorte industrial externa son que se reduce el sesgo debido al efecto del trabajador sano y que las dos poblaciones probablemente serán muy similares en cuanto a factores relacionados con el estilo de vida. Por tanto, la importancia de factores potencialmente confusores, como el tabaquismo, será tal vez mínima. Sin embargo, entre los trabajadores manuales puede haber considerables diferencias en los hábitos de vida dependiendo del grado de educación, el tipo de trabajo, la experiencia profesional, etc.<sup>17</sup>. Además, la población de referencia podría estar también expuesta a factores de riesgo para la enfermedad que se está investigando. Por ejemplo, en un estudio danés con trabajadores expuestos a herbicidas, en el que se investigaba la incidencia de linfomas no hodgkinianos, se encontró que la incidencia de linfomas era más alta en el grupo de referencia que en el grupo expuesto y en la población general, debido probablemente a la presencia de algún factor ocupacional no identificado entre los trabajadores de la cohorte de referencia<sup>18</sup>.

Cuando se utiliza un grupo interno de referencia, las comparaciones se restringen a la experiencia de las personas-tiempo de la propia cohorte. En este caso, la categoría de referencia consiste en las tasas de enfermedad asociadas a las personas-tiempo de los individuos no expuestos o poco expuestos. En este tipo de estudios, tanto la cantidad como la calidad de los datos son similares en los grupos de comparación. Los sesgos de selección (incluyendo el efecto del trabajador sano) tenderán a ser similares entre los individuos expuestos y los no expuestos. Sin embargo, tasas inestables en el grupo de referencia pueden limitar la precisión de los riesgos estimados, especialmente cuando se investigan enfermedades poco comunes. Además, la variabilidad de la exposición puede ser pequeña dentro de la cohorte y, por tanto, puede que no sea posible identificar a un grupo interno de referencia de individuos no expuestos.

#### Comparación de tasas: el cálculo de la razón de mortalidad estandarizada

La comparación de tasas específicas entre dos poblaciones puede resultar especialmente incómoda si tenemos en cuenta el gran número de estratos edad-período existentes para cada una. Por tanto, es necesario utilizar una medida que resuma toda esta información específica para cada población y nos permita comparar fácilmente cada una ajustando por las diferencias estructurales de cada población según la edad y el período. La estandarización es un método estadístico que nos permite realizar estas comparaciones controlando según los efectos de posibles fac-

tores de confusión. Existen dos métodos de estandarización, directo e indirecto, que difieren esencialmente en el tipo de peso utilizado por el cálculo de las tasas estandarizadas.

El método indirecto de estandarización es el más frecuentemente aplicado en cohortes laborales, ya que no requiere el conocimiento de las tasas específicas de mortalidad de la cohorte, sino que sólo es preciso conocer el número total de muertes (casos) ocurridos y la distribución de personas-años por edad y período de tiempo de la cohorte. Mediante este método podemos calcular el número de muertes esperadas en la cohorte de estudio si ésta tuviera las mismas tasas específicas de mortalidad que la población de referencia que, al contrario de la cohorte, sí que deben de ser conocidas. Multiplicando las tasas de mortalidad específicas de la población de referencia por el término «personas-tiempo a riesgo» en el correspondiente estrato edad-tiempo obtenemos el número de muertes esperadas en la cohorte. El cociente entre el número de muertes observadas y esperadas se conoce como RME, y puede definirse como un media ponderada de las razones de tasas específicas edad-período<sup>19</sup>, donde los pesos son el número esperado de muertes en la cohorte. Algunas veces se expresa la RME como un porcentaje, aunque esto no es necesario y es preferible expresar según la unidad, al igual que otros estimadores del riesgo. Definiremos la RME tal y como se expresa en la tabla 1.

#### Cuadro 5. Cálculo del número de muertes esperadas en la cohorte de plomistas

Para el cálculo del número de casos esperados es preciso conocer las tasas específicas de la población de referencia. En este caso, han sido extraídas del libro «Epidemiología y Prevención del Cáncer en Cataluña 1975-1992» publicado por el Servei Català de la Salut de la Generalitat de Catalunya<sup>20</sup>. Las tasas de mortalidad disponibles en este registro corresponden al período 1975-1992. Puesto que el período de seguimiento de la cohorte abarca el período 1964-1999, se han tenido que estimar las tasas de mortalidad pertenecientes a 1964-1974 y 1993-1999. Se ha aplicado un modelo de extrapolación teniendo en cuenta que la tendencia de la mortalidad ha presentado un incremento lineal anual del 2,1% para todos los cánceres

Finalmente, se evalúa la magnitud del estimador de riesgo RME y también la significación estadística. Una RME superior a 1 implica que existe un riesgo de mortalidad superior entre la cohorte estudiada en comparación con la población de referencia. Los intervalos de confianza del 95% de dichas estimaciones se pueden calcular siguiendo diferentes fórmulas<sup>21</sup> y normalmente dan como salida rutinaria los paquetes estadísticos. Una RME se considera estadísticamente significativa si su correspondiente intervalo de confianza (calculado a partir de un nivel de significación del 5%) no comprende la unidad.

Cuadro 6. Riesgo de morir por cáncer en la cohorte de plomistas

A partir de las estimaciones de RME (riesgo en comparación con la población masculina de Cataluña, ajustado por edad y período de tiempo) y de los respectivos intervalos de confianza del 95%, se ha observado una mortalidad prácticamente el doble de lo esperada para todos los cánceres (RME = 1,96). Sin embargo, este riesgo no es estadísticamente significativo. También se ha observado un incremento del riesgo por cáncer de pulmón de casi 5 veces superior a lo esperado en comparación al riesgo entre la población de Cataluña, y esta diferencia sí es estadísticamente significativa (RME = 4,79)

	Personas- años	Muertes observadas	Muertes esperadas	RME	IC del 95%
<b>Cáncer total</b>					
Total	910,11	4	2,03	1,96	(0,73-5,24)
Empleo	357,92	1	0,57	1,74	(0,24-12,4)
<10 años					
Empleo	552,19	3	1,46	2,05	(0,66-6,36)
>10 años					
Latencia de	546,48	4	1,85	2,15	(0,81-5,74)
10 años					
Latencia de	255,17	3	1,37	2,18	(0,7-6,78)
20 años					
<b>Cáncer de pulmón</b>					
Total	910,11	3	0,63	4,79	(1,54-14,8)
Empleo	357,92	0	0,18	0	-
<10 años					
Empleo	552,19	3	0,45	6,65	(2,14-20,6)
>10 años					
Latencia de	546,48	3	0,58	5,13	(1,65-15,9)
10 años					
Latencia de	255,17	3	0,44	6,78	(2,18-21,0)
20 años					

Análisis de variables relacionadas con el tiempo (latencia, duración del empleo)

Muchas de las enfermedades de interés en estudios laborales son enfermedades crónicas, que ocurren muchos años después de la primera exposición. Los análisis epidemiológicos que examinan relaciones dosis-respuesta deben tener en cuenta el período de latencia de la enfermedad de estudio y deben evaluar qué períodos de exposición fueron los más relevantes para la inducción o progresión de la enfermedad. Si la enfermedad de interés es, por ejemplo, una enfermedad neoplásica, cualquier exposición a un carcinógeno industrial poco antes del diagnóstico será totalmente irrelevante para el desarrollo de la enfermedad y debe ser ignorado al calcular la dosis de exposición. Igualmente, cada neoplasia que aparezca justo antes de la exposición no se puede atribuir a dicha exposición. El término general que recibe este procedimiento es el de «análisis de latencia o método con retardo»<sup>22, 23</sup>. Todos los análisis de

latencia comprenden alguna forma de ponderación de la exposición, de modo que las exposiciones asumidas etiológicamente importantes son ponderadas en mayor magnitud que los períodos presumiblemente irrelevantes. Las variables relacionadas con el tiempo pueden actuar directa o indirectamente como indicadores de intensidad de exposición (p. ej., duración del empleo), posibles factores de confusión entre exposición y riesgo de enfermedad (p. ej., tiempo desde inicio exposición) e incluso como modificadores del efecto (p. ej., edad inicio empleo o tiempo desde la última exposición). La superposición entre el período de exposición y el seguimiento y, por otro lado, el carácter crónico de algunas enfermedades influyen directamente en la relación entre exposición y respuesta.

La aplicación de períodos ventana (time windows) constituye una extensión del modelo con retardo<sup>24</sup>. Este procedimiento se basa esencialmente en identificar el período de exposición específico más relevante para la ocurrencia de la enfermedad. En el método de períodos ventana no se tienen en cuenta las exposiciones ocurridas en períodos tempranos ni las ocurridas en períodos tardíos. La identificación del período de exposición más apropiado se realiza estimando el riesgo en diferentes ventanas de tiempo consecutivas. Estos métodos se han desarrollado principalmente en epidemiología laboral del cáncer y se basan en la asunción de un modelo poli-factorial y de múltiples etapas de causalidad de la enfermedad.

Cuadro 7. Resultados por latencia y duración en la cohorte de plomistas

En la cohorte de plomistas se estimaron también las RME por tiempo de duración de empleo (menos de 10 años y más de 10 años de empleo) y por latencia (10 o más años, y 20 o más años). La mortalidad por todos los cánceres aumenta ligeramente con la latencia y con la duración de la exposición (cuadro 6). La mortalidad por cáncer del pulmón aumenta con la latencia, hasta casi 7 veces lo esperado en un período de 20 años desde la primera exposición. También se ha observado un exceso de riesgo estadísticamente significativo de 6,65 para el grupo de trabajadores de más de 10 años de ocupación en la empresa. Estos resultados indican que el cáncer del pulmón está asociado con una exposición laboral. Se debe tener en cuenta que las variables relacionadas con el tiempo (duración, latencia, pero también edad y período de tiempo) pueden estar correlacionadas. Así pues, el aumento del riesgo entre plomistas con más de 10 años de empleo puede ser debido a que éstos, que también han empezado a trabajar más pronto en la empresa, son los que entran también en la categoría de más de 20 años de latencia. La evaluación de la importancia de cada variable temporal se puede hacer con un análisis estratificado, calculando la RME para las diferentes categorías de latencia por duración, o también aplicando métodos multivariados. Una u otra solución se puede aplicar solamente si las cohortes son grandes y hay suficientes efectivos de personas-años acumulados en cada estrato

## Evaluación de los factores de confusión

En los estudios epidemiológicos la confusión se puede prevenir en la fase de diseño del estudio o se puede controlar en el análisis. El control de la confusión en la fase del diseño del estudio se realiza durante la selección de la población de referencia, por ejemplo, apareando los individuos expuestos y no expuestos en variables como edad, sexo o clase social. En la fase de análisis, la evaluación de la confusión se realiza aplicando el análisis estratificado o métodos de análisis multivariados. En este caso se comparan las estimaciones crudas del efecto con las estimaciones obtenidas después de eliminar la confusión. Un requisito indispensable es la disponibilidad de información acerca de la exposición al factor confusor. En la mayoría de estudios laborales, y especialmente en los estudios históricos de cohortes, raramente se dispone de información sobre factores potencialmente confusores. Por tanto, con frecuencia es necesario evaluar indirectamente el efecto de un factor confusor potencial.

### Estimación indirecta de la magnitud del efecto de la confusión

En epidemiología laboral, los factores de riesgo generales de enfermedad, como el tabaquismo o la obesidad, se pueden considerar como explicaciones alternativas de un exceso de mortalidad o morbilidad asociado a una exposición laboral. Estas explicaciones alternativas son plausibles ya que la prevalencia de factores relacionados con el estilo de vida, como el hábito tabáquico, varía dentro de una misma población y está frecuentemente asociada al estado socioeconómico. En España, al igual que en muchas sociedades occidentales, la frecuencia de tabaquismo es mayor en los trabajadores manuales del sexo masculino que en los trabajadores no manuales, aunque ello no ocurre necesariamente en las mujeres<sup>25</sup>. Frecuentemente, no se dispone de información sobre la prevalencia de factores de riesgo generales, como el tabaco o el alcohol, sobre todo en los estudios históricos de cohortes. En estos casos, la evaluación de la magnitud del efecto de los factores de confusión sólo se puede realizar de forma indirecta.

Uno de los métodos indirectos más sencillos ha sido propuesto por Axelson<sup>26</sup>. Con este método se obtiene una estimación aproximada de la magnitud del efecto confusor de factores de riesgo específicos. Sin embargo, para ello es necesario estimar la prevalencia del factor de confusión en la población de estudio (o en una muestra de la misma), así como en la población de referencia. Además, es necesario estimar la magnitud del riesgo relativo asociado a la exposición al factor del confusión. El efecto del factor de confusión en el estimador del riesgo de enfermedad se calcula separadamente en los individuos expuestos y en los no expuestos al factor de interés, en relación con las tasas de los sujetos no expuestos al factor de confusión. Entonces, se deriva un estimador relativo del grado de riesgo entre los grupos de exposición que puede ser ex-

plicado por el factor de confusión. Este método asume que no existe interacción entre la exposición en estudio y el factor de confusión.

En el ejemplo que sigue se estima el efecto confusor potencial del tabaquismo en un estudio sobre cáncer de pulmón en una población expuesta a un carcinógeno laboral. En la población de referencia (nacional) la prevalencia de fumadores es del 50% ( $p_f = 0,5$ ), siendo el 50% restante correspondiente a la población de no fumadores ( $p_{nf} = 0,5$ ). En la población expuesta al factor de riesgo laboral la prevalencia de tabaquismo es del 60% ( $p_f = 0,6$ ), siendo el 40% restante no fumadores ( $p_{nf} = 0,4$ ). El riesgo relativo estimado de mortalidad por cáncer de pulmón en las dos poblaciones es de 1 en los no fumadores y de 10 en los fumadores. La tasa global de incidencia en cada grupo de exposición puede ser considerada como una media ponderada de la incidencia en fumadores y no fumadores. Así, en los individuos no expuestos la incidencia sería:  $I_{ne} = I_{nf}(p_{nf}) + I_f(p_f)$ . Una ecuación similar se puede derivar para los individuos expuestos:  $I_e = I_e(p_{nf}) + I_f(p_f)$ . En este ejemplo concreto, la incidencia en la población de referencia es:  $I_{ne} = I_{nf}(0,5) + I_f(10)(0,5) = 5,5$  ( $I_{nf}$ ). O sea, la incidencia global en el grupo de no expuestos (población de referencia) es 5,5 veces más alta que la incidencia en los no fumadores de esta población. La incidencia en los expuestos es  $I_e = I_{nf}(0,4) + I_f(10)(0,6) = 6,4$  ( $I_{nf}$ ). Esto es, la incidencia global en la población expuesta es 6,4 veces mayor que la incidencia de los no fumadores en dicha población. Por tanto, se puede calcular una razón que estime el efecto específico del factor confusor. Esta estimación se puede utilizar para ajustar el riesgo observado de exposición o, simplemente, como una guía para evaluar la importancia del efecto confusor potencial del tabaco. En este ejemplo, se puede estimar que el hábito tabáquico produce un aumento de la incidencia de cáncer de pulmón de aproximadamente el 15% en los expuestos respecto a los no expuestos ( $6,4/5,5 = 1,15$ ). Por tanto, cualquier exceso de riesgo por encima del 15% se pueda atribuir, muy probablemente, a la exposición laboral de interés.

Existen otros métodos para evaluar el efecto de factores confusores potenciales. Uno de ellos consiste en comparar la incidencia de enfermedades con factores de riesgo comunes y derivar una estimación indirecta de la importancia de la exposición al factor confusor. Por ejemplo, la incidencia de diversas enfermedades (neoplásicas o no) se ha asociado al tabaquismo. De este modo, si en un estudio se observa un riesgo alto de cáncer de pulmón, pero no se observa un exceso de riesgo para otras enfermedades relacionadas con el tabaco (cáncer de vejiga, laringe, esófago, u otras enfermedades respiratorias), es muy probable que el riesgo observado de cáncer de pulmón en la población de estudio no sea únicamente debido a una mayor prevalencia de tabaquismo, sino a la exposición al factor laboral de interés y/o a otra exposición desconocida.

Otro procedimiento es realizar equiparaciones internas dentro de la misma cohorte, comparando los individuos expuestos con los no expuestos o con los trabajadores con niveles bajos de exposición, asumiendo que la distribu-

**Cuadro 8. Evaluación indirecta del posible efecto confusor del tabaquismo en la cohorte de plomistas utilizando el método propuesto por Axelson**

El consumo de tabaco es un factor de riesgo muy importante para el cáncer del pulmón. Se podía asumir que la proporción de fumadores entre los plomistas es más alta que dicha proporción en la población general. Actualmente, no existen registros de mortalidad por cáncer entre fumadores y no fumadores en España. Se ha evaluado indirectamente la posible sobrestimación del riesgo producido por el efecto confusor del tabaco, basado en las diferencias de consumo de tabaco entre la clase trabajadora y la población general masculina de Cataluña<sup>27</sup>. Así, a partir de las entrevistas con los plomistas, se ha calculado que la prevalencia del tabaquismo variaba entre un 70 y un 90%. Asimismo, en Cataluña se ha asumido que la proporción del tabaquismo entre varones era de 50%. Las estimaciones de riesgo sobre tabaco y cáncer y tabaco y cáncer del pulmón utilizadas han sido las del estudio de los médicos ingleses<sup>28</sup>.

Como puede observarse a continuación, la más alta prevalencia del tabaquismo entre los plomistas (90%) podía producir solamente un aumento del riesgo del orden de un 30% para todos los cánceres, y del orden de un 70% para el cáncer de pulmón. Si comparamos estos resultados con el aumento del riesgo obtenido para la cohorte, constatamos que el exceso de cáncer, y especialmente el de pulmón observado entre los plomistas, es mayor que el que puede explicarse por las diferencias de hábitos tabáquicos. Por tanto, aunque el tabaquismo indudablemente ha contribuido al aumento del riesgo entre los plomistas, éste no se puede atribuir solamente al tabaco, por lo que tiene que estar asociado con exposiciones laborales

Tipos de cáncer	Riesgo identificado	Riesgo que se podría explicar por el tabaco según el porcentaje de fumadores en la cohorte		
		90%	70%	50%
Todos los cánceres	1,96	1,30	1,15	1
Cáncer de pulmón	4,79	1,70	1,35	1

ción del factor de confusión es la misma entre los dos grupos. Esta asunción es frecuentemente razonable cuando se comparan trabajadores manuales empleados en diferentes departamentos de una misma industria.

**Conclusiones**

El objetivo principal de los estudios de cohortes laborales es medir la relación entre la exposición laboral y la incidencia de la enfermedad o la mortalidad. El cálculo de la RME se encuentra entre los métodos de análisis más frecuentemente aplicados para evaluar el riesgo asociado a una exposición laboral. Para su aplicación, normalmente se utilizan grupos de comparación externos, que la mayoría de las veces suelen ser las tasas nacionales de mortalidad. Por otro lado, existe una gran variedad de programas y paquetes estadísticos que permiten realizar este tipo de análisis con relativa facilidad. En estos últimos años, y

de manera generalizada, han aumentado las fuentes de datos que permiten obtener información sobre registros de mortalidad, ya sea a través de publicaciones o mediante consultas en Internet, en páginas pertenecientes a organismos e instituciones vinculadas a la investigación médica.

En el análisis de datos prospectivos es necesario calcular el dato de las personas-años a riesgo que corresponde a la experiencia de la población de estudio durante la cual los sujetos están a riesgo de padecer la enfermedad. Cuando se evalúa la relación entre exposición y enfermedad es importante tener en cuenta el período de latencia de cada enfermedad y proporcionar estimadores de riesgo (RME), tanto para todo el período evaluado como para «ventanas de tiempo» que corresponden al período crítico de exposición que, supuestamente, han provocado la enfermedad. Cuando se hacen comparaciones con poblaciones externas es importante tener en cuenta los posibles factores de confusión, como el tabaco. Sin embargo, en muchos estudios laborales dicha información no está disponible. En estos casos, se pueden aplicar otros métodos indirectos para evaluar la magnitud aproximada del efecto de posibles factores de confusión.

**BIBLIOGRAFÍA**

1. Hammond EC, Selikoff IJ, Seidman H. Asbestos exposure, cigarette smoking and death rates. *Ann NY Acad Sci* 1979; 330: 473-490.
2. Case RAM, Hosker ME, McDonald DB, Pearson JD. Tumors of the urinary bladder in workmen engaged in the manufacture and use of certain dyestuff intermediates in the British chemical industry. *Br J Med* 1954; 11: 75-104.
3. Kogevinas M, Becher H, Benn T, Bertazzi PA, Boffetta P, Bueno de Mesquita HB et al. Cancer mortality in workers exposed to phenoxy herbicides, chlorophenols and dioxins. An expanded and updated international cohort study. *Am J Epidemiol* 1997; 145: 1061-1075.
4. Sala Serra M, Sunyer J, Kogevinas M, McFarlane D, Antó JM. Cohort study on cancer mortality among workers in the pulp and paper industry in Catalonia, Spain. *Am J Ind Med* 1996; 30: 87-92.
5. Rodríguez Artalejo F, Castano Lara S, De Anrés Manzano B, García Ferruelo M, Iglesias Martín L, Calero JR. Occupational exposure to ionising radiation and mortality among workers of the former Spanish Nuclear Energy Board. *Occup Environ Med* 1997; 54: 202-208.
6. Checkoway H, Pearce N, Dement JM. Design and conduct of occupational epidemiology studies (II). Analysis of Cohort Data. *Am J Ind Med* 1989; 15: 375-394.
7. Dos Santos Silva I. *Cancer epidemiology: principles and methods*. Lyon: International Agency for Research on Cancer, 1999.
8. PAMCOMP 1.10 for Windows 98/95/NT (programa informático). University of Münster (Alemania): Institute of Epidemiology and Social Medicine, 1999. Disponible en: <http://medweb.uni-muenster.de/institute/epi/pamcomp/pamcomp.html>.
9. OCMAP-PLUS 2.0 for Windows (programa informático). University of Pittsburgh (EE.UU.): Department of Biostatistics Graduate School of Public Health, 1999. Disponible en: <http://ocmap.biostat.pitt.edu/ocmore.html>.

10. PC-LTAS 1.0c (programa informático). Ohio (EE.UU.): US Department of Health and Human Services. National Institute for Occupational Safety and Health, 2000. Disponible en: <http://www.cdc.gov/niosh/ltindex.html>.
11. Intercooled STATA 6.0 for Windows 98/95/NT (programa informático). Stata Corporation. Texas (EE.UU.), 1999.
12. Instituto Nacional de Estadística (INE), Madrid, 2000. Disponible en: <http://www.ine.es/daco42/sanitarias/dacocem.htm>.
13. WHO Statistical Information System (WHOSIS). Ginebra: World Health Organization (OMS), 2000. Disponible en: <http://www.who.int/whosis/>.
14. International Classification of disease. Manual of the International Statistical Classification of Diseases, Injuries and Causes of Death (10.<sup>a</sup> revisión). Ginebra: World Health Organization, 1987. Disponible en: <http://www.who.int/whosis/icd10/electron.htm>.
15. Parkin DM, Whelan SL, Ferlay J, Raymond L, Young J. Cancer incidence in five continents (Vol VII). Lyon: IARC Scientific Publications Number 143, 1997. Disponible en: <http://www-dep.iarc.fr/resour/software/ci5V7sof.htm>.
16. López-Abente G, Pollán M, Escolar A, Errezola M, Abaira V. Atlas de mortalidad por cáncer y otras causas en España 1978-1992. Disponible en: [www2.nca.es/hospital/atlas92/www/Atlas92.html](http://www2.nca.es/hospital/atlas92/www/Atlas92.html)
17. Rodríguez V, Tardon A, Kogevinas M, Prieto CS, Cueto A, García M et al. Lung cancer risk in iron and steel foundry workers: a nested case control study in Asturias, Spain. *Am J Ind Med* 2000; 28: 644-650.
18. Lynge E. Cancer in phenoxy herbicide manufacturing workers in Denmark, 1947-87 —an update. *Cancer Causes Control* 1993; 4: 261-272.
19. Rothman KJ, Greenland S. *Modern Epidemiology* (2.<sup>a</sup> ed.). Filadelfia: Lippincott-Raven Publishers, 1998.
20. Borràs JM, Borràs J, Viladiu P, Bosch FX. *Epidemiología y prevención del cáncer en Catalunya (1975-1992)*. Barcelona: Institut Català d'Oncologia, 1997.
21. Breslow NE, Day NE. *Statistical methods in cancer research (vol. II). The design and analysis of cohort studies*. Lyon: IARC Scientific Publication N.º 82, 1987.
22. Checkoway H, Pearce NE, Crawford-Brown DJ. *Research methods in occupational epidemiology*. Nueva York: Oxford University Press, 1989.
23. Pearce N. Methodological problems of time-related variables in occupational cohort studies. *Rev Epidem Sante Publ* 1992; 40: 543-554.
24. Pearce N. Multistage modelling of lung cancer mortality in asbestos textile workers. *Int J Epidemiol* 1988; 17: 747-752.
25. Regidor E, Gutiérrez-Fisac J, Rodríguez C. *Diferencias y desigualdades en salud en España*. Madrid: Díaz de Santos, 1994.
26. Axelson O. Aspects on confounding in occupational health epidemiology. *Scand J Work Environ Health* 1978; 4: 98-102.
27. Faggiano F, Partanen T, Kogevinas M, Boffetta P. Socio-economic differences in cancer incidence and mortality. En: Kogevinas M, Pearce N, Susser M, Boffetta P, editores. *Social inequalities in cancer*. Lyon: IARC Scientific Publications N.º 138, 1997; 65-176.
28. Doll R, Peto R, Wheatley K, Gray R, Sutherland I. Mortality in relation to smoking: 40 years' observations on male British doctors. *BMJ* 1994; 309: 901-911.