

DIAGNÓSTICO ESTÁNDAR PARA VALIDAR LAS CAUSAS DE MUERTE CERTIFICADAS

Fernando García Benavides * / Vicente Arraez ** / Andreu Nolasco * / Lucas Jiménez *** / Pascual Bordes *** / Francisco Bolumar *

(*) Departamento de Salud Comunitaria. Universidad de Alicante.

(**) Hospital General de Elche. INSALUD. Alicante.

(***) Hospital del INSALUD. Alicante.

Resumen

Analizamos la concordancia entre dos médicos, los cuales a partir de idéntica información clínica correspondiente a 1.153 defunciones, establecen sus respectivas causas de muerte. El índice Kappa para el total de causas es de 0.78, oscilando entre 0.66, enfermedades de sistema nervioso, y 0.96, tumores. Concluimos que estas causas de muerte así establecidas pueden servir como diagnóstico estándar para validar las causas de muerte certificadas.

Palabras clave: Concordancia. Causas de Muerte.

STANDARD DIAGNOSTIC TO VALIDATE CERTIFIED CAUSES OF DEATH

Summary

We analyse the agreement between two physicians who establish cause of death for 1,153 cases using the same clinical information. The Kappa index for the total of causes is 0.78, varying between 0.66 for diseases of the nervous system and 0.96 for tumours. We conclude that these causes of death thus established can serve as a standard diagnosis to validate certified causes of death.

Key words: Concordance. Causes of death.

Introducción

Para valorar la exactitud de una medición se precisa de un modelo de referencia independiente que pueda aceptarse como fiable y frente al cual comparemos nuestras observaciones y resultados, sean éstos del tipo que fueren: observaciones clínicas, tests diagnósticos, etc. En el presente trabajo el objetivo a validar lo constituyen las causas de muerte certificadas.

En algunos casos el test estándar o modelo de referencia es relativamente accesible y poco costoso. Sin embargo, esto no es lo frecuente, es más, en algunas ocasiones se dispone teóricamente de un buen estándar, pero por razones asociadas al coste u otros problemas -aceptación social, complejidad técnica o problemas éticos- no pueden ser utilizados para el propósito señalado¹.

Este es nuestro caso, nosotros disponemos teóricamente de los diagnósticos que suministran las autopsias como el mejor estándar posible para conocer con qué exactitud se

certifican las causas de las defunciones. Sin embargo partiendo de la sospecha, más tarde confirmada, de que el porcentaje de autopsias en nuestro país es aún muy bajo, tuvimos que construir un estándar alternativo a partir de la información clínica recogida en diversas fuentes: historias clínicas, cuestionarios e informaciones forenses².

El inconveniente de este tipo de estándar es que siempre queda la duda de su propia inexactitud, ya que no ofrece garantías por sí mismo de estar midiendo aquello que deseamos medir.

Una manera de comprobar la fiabilidad del estándar así construido consiste en hacer intervenir a más de un tasador, los cuales, sobre la misma información han de establecer sus propios resultados. De este modo, si la concordancia entre ellos es alta nos indica la posibilidad de que la medición realizada refleja aquello que deseamos medir, y por lo tanto podemos utilizarlo con garantías para validar, en este caso, las causas de muerte certificadas.

Material y método

A partir de la información clínica disponible para 1.176 defunciones, cuya recogida y el estudio de sus características ha sido objeto de otro trabajo², dos médicos, especialista en medicina intensiva el primero y en cardiología el segundo, cumplieron, cada uno de ellos y de manera independiente, un nuevo certificado de defunción para 1.153 de aquéllas. Hubo 23 defunciones en las que sólo intervino un médico, por lo que no se tienen en cuenta para el presente trabajo.

Con el fin de asegurar que ambos médicos conocían y utilizaban los mismos criterios de clasificación, se sometieron a un ensayo con 30 definiciones escogidas al azar, tras el cual fueron discutidas las distintas causas de muerte así como su certificación³. Ambos médicos utilizaron a lo largo del estudio el mismo documento de notificación, el modelo de certificación de causas de muerte del Instituto Nacional de Estadística.

Consideramos, en una primera etapa, la variabilidad a nivel de aquellos

Consideramos, en una primera etapa, la variabilidad a nivel de aquellos grupos de causas de mayor importancia cuantitativa, dejando una categoría residual donde se incluyen los grupos de enfermedades de escasa relevancia en nuestra muestra. Incluimos otra categoría, denominada no útil, donde se clasifican aquellas defunciones para las cuales la información clínica disponible fue valorada, por parte de uno o de los dos médicos clasificadores, como no útil para establecer un diagnóstico de muerte. En una segunda etapa, estudiamos la concordancia a nivel de entidades nosológicas más específicas.

Para expresar el grado de concordancia hemos optado por una escala binaria: concuerdan con la categoría diagnosticada o no concuerdan, considerando dentro de esta segunda opción a todas las demás categorías diagnosticadas.

El índice utilizado para resumir la

concordancia ha sido el índice Kappa, ya que es, de entre todos los propuestos⁴ el que permite conocer la variabilidad entre observadores una vez eliminada la concordancia debida únicamente al azar, $K = (p_o - p_e) / (1 - p_e)$ siendo p_o la proporción de concordancia observada y p_e la proporción de concordancia esperada. Un test de normalidad Z se aplicó para comprobar si el índice Kappa obtenido era estadísticamente diferente de cero.

Los valores de este índice varían de -1 a +1, donde +1 significa concordancia perfecta, cero concordancia debida al azar y -1 discordancia completa.

Por último, analizamos mediante el índice de McNemar $(b-c)/(b+c)$ el patrón de discordancia entre los dos médicos clasificadores, representando los valores próximos a 0 que la discordancia no está sesgada hacia ninguno de los observadores, y los valores próximos a +1 o -1 que uno

de los médicos clasificadores tiende a clasificar hacia una categoría, i.e. cardiovasculares, más que en las otras o al contrario. Con el test de $X_m^2 = (b-c)^2 / (b+c)$ comprobamos si las diferencias encontradas son estadísticamente significativas⁵.

Resultados

En la tabla 1, donde se presenta la matriz de concordancia para aquellos grupos de causas de mayor importancia cuantitativa, aparecen en las columnas las frecuencias de diagnósticos establecidos por el médico certificador 1 y en las filas aquellos correspondientes al médico certificador 2. Así, por ejemplo, mientras el primero opina que 59 defunciones han sido debidas a enfermedades endocrinas, el médico certificador 2 piensa que 12 de estas son consecuencia de enfermedades del aparato

Tabla 1. Matriz de concordancia para los grandes grupos de causas de muerte. (Frecuencias)

	CIE-9 ^o	MÉDICO CERTIFICADOR 1												TOTAL	
		INFECC.	TUM.	ENDCR.	NERV.	CIR.	RESP.	DIG.	G-UR.	ANOM.	EXT.	OTRAS.	NO/UT.		
MÉDICO CERTIFICADOR 2															
INFECCIOSAS	(001-139)	7	0	0	0	0	0	0	0	0	1	0	0	8	
TUMORES	(140-239)	0	285	1	1	3	0	1	0	0	1	0	0	292	
ENDOCRINAS	(240-279)	0	0	40	0	6	0	0	0	0	0	0	0	46	
NERVIOSAS	(320-389)	0	0	0	6	1	1	1	0	0	0	0	0	9	
CIRCULATORIO	(390-459)	0	2	12	2	295	12	1	4	0	2	1	2	333	
RESPIRATORIO	(460-519)	0	1	0	0	5	60	0	1	1	2	1	1	72	
DIGESTIVO	(520-579)	1	1	0	0	0	0	84	0	0	1	0	0	87	
GENITOURINARIO	(580-629)	0	0	1	0	5	0	0	23	0	0	0	1	30	
ANOMALÍAS	(740-759)	0	0	0	0	0	0	0	0	6	0	0	0	6	
EXTERNAS	(800-999)	0	0	0	0	1	1	0	1	0	71	1	0	75	
OTRAS	-	0	0	0	0	4	0	1	0	0	1	9	0	15	
NO ÚTIL	-	1	4	5	0	75	9	6	8	0	5	1	66	180	
TOTAL	-	9	293	59	9	395	83	94	37	7	84	13	70	1153	

Tabla 2. Concordancia entre dos médicos al certificar las causas de muerte

GRUPOS DE CAUSAS	CIE 9 ^a	KAPPA	INTERV.CONF.	Z
INFECCIOSAS	(001-139)	0,82	0,77-0,87	28,27*
TUMORES	(140-239)	0,96	0,91-1,00	33,10*
ENDOCRINAS	(240-279)	0,75	0,69-0,80	25,86*
S.NERVIOSO	(320-389)	0,66	0,61-0,71	22,75*
CIRCULATORIO	(390-459)	0,72	0,67-0,77	24,82*
RESPIRATORIO	(460-519)	0,75	0,70-0,80	25,86*
DIGESTIVO	(520-579)	0,92	0,87-0,97	31,72*
GENITOURINARIO	(580-629)	0,67	0,62-0,72	23,10*
ANOMALÍAS CONGEN.	(740-759)	0,92	0,87-0,97	31,72*
CAUSAS EXTERNAS	(800-999)	0,88	0,83-0,93	30,34*
OTRAS	-	0,85	0,80-0,90	29,31*
NO ÚTIL	-	0,48	0,46-0,51	19,20*
TOTAL	-	0,78	0,76-0,81	60,33*

* P < 0.001

circulatorio. El médico clasificador 2 incluye un mayor número de defunciones que el médico certificador 1, en la categoría no útil 180 frente a 70 respectivamente.

Los índices Kappa de estos mismos grupos de causas, tal como vemos en la tabla 2, muestran una concordancia casi perfecta, entre 1 y 0,80, para las infecciones, tumores, enfermedades digestivas, anomalías congénitas y causas externas; una concordancia aceptable, entre 0,80 y 0,60 para el resto de grupos de causas consideradas, destacando el índice Kappa obtenido para el total, 0,78. En la categoría no útil encontramos un índice de 0,48, que puede ser catalogado como de una concordancia moderada.

Sin embargo, al analizar estadísticamente los índices observados, comprobamos que para todos ellos, incluido el de la categoría no útil, existen diferencias estadísticamente significativas ($p < 0,001$).

Para entidades nosológicas más específicas, el índice Kappa, se sitúa

también por encima de 0,60, excepto para las enfermedades agrupadas en la rúbrica «cardiopatía isquémica» (411-414), aunque el límite superior de su intervalo de confianza, 0,63, está dentro del rango de 0,60-0,80 considerado como concordancia aceptable (tabla 3).

Al analizar los patrones de discordancia entre ambos certificadores (tabla 4), encontramos que los índices de McNemar que muestran sesgos estadísticamente significativos son los que se refieren a la categoría de enfermedades endocrinas, -0,52, enfermedades circulatorias, -0,45, causas externas, -0,53 y No útil, 0,93. Este último índice debe interpretarse en distinto sentido que el resto, pues mientras signo negativo indica que el médico revisador 1 tiende a clasificar dentro de estas categorías (endocrinas, circulatorias y causas externas) un mayor número de defunciones que el médico certificador 2, el índice de McNemar referido a la categoría de no útil muestra que el médico certificador 2 clasifica dentro de esta cate-

goría mayor número de defunciones que el médico certificador 1, siendo estas diferencias estadísticamente significativas.

Si volvemos a la tabla 1, encontramos que, efectivamente, de las 395 defunciones que el médico certificador 1 clasifica como circulatorias, hay 100 donde no coinciden este diagnóstico con el dado por el segundo observador; mientras que este último, de las 333 defunciones que clasifica como circulatorias, no coinciden con el médico certificador 1 en 38; destacando que de aquellas 100 primeras, 75 son clasificadas por el médico clasificador 2 como no útiles.

Discusión

Estos resultados pueden ser analizados en una doble dirección; por un lado son suficientemente ilustrativos como para hacernos pensar que, una vez constatada la alta consistencia entre ambos médicos observadores, la causa de muerte obtenida mediante la vía propuesta, puede servir como

Tabla 3. Concordancia entre dos médicos al certificar las causas de muerte

GRUPOS DE CAUSAS	CIE 9ª	KAPPA	INT.DE CONF.	Z
T.M. DIGESTIVOS	(150-159)	0,94	0,88-0,99	32,41*
T.M. RESPIRATORIOS	(161-165)	0,97	0,92-1,00	33,44*
INF.AG. MIOCARD.	(410)	0,77	0,72-0,82	26,55*
OTRAS CARD. ISQUEM.	(411-414)	0,58	0,53-0,63	20,00*
ENF. CEREBROVASC.	(430-438)	0,79	0,74-0,84	27,24*
CIRROSIS HEPÁTICA	(571)	0,96	0,91-1,00	33,10*

* $P < 0,001$

Tabla 4. Patrones de discordancia entre dos médicos al certificar las causas de muerte

GRUPOS DE CAUSAS	CIE 9ª	McNEMAR	CHI-CUADRADO
INFECCIOSAS	(001-139)	-0,33	0,33
TUMORES	(140-239)	-0,06	0,06
ENDOCRINAS	(240-279)	-0,52	6,76**
S.NERVIOSO	(320-389)	0,00	0,00
CIRCULATORIO	(390-459)	-0,45	27,85**
RESPIRATORIO	(460-519)	-0,31	3,45
DIGESTIVO	(520-579)	-0,53	3,76
GENITOURINARIO	(580-629)	-0,33	2,33
ANOM. CONGENITAS	(740-759)	-1,00	1,00
CAUSAS EXTERNAS	(800-999)	-0,53	4,76*
OTRAS	-	0,20	0,40
NO ÚTIL	-	0,93	102,54**

* $p < 0,05$

** $p < 0,001$

diagnóstico estándar frente al diagnóstico seleccionado del certificado original, validando con ello su exactitud. Esta propuesta ha de ser completada con la intervención de un tercer observador que actúe en los casos de discordancia, dando la razón a uno de los primeros.

Por otro lado, interesa no solo constatar que existe concordancia y cual es su magnitud, sino también buscar las fuentes de desacuerdo, pues si bien ésta es pequeña en la mayoría de los grupos de causas analizados, puede ser de gran ayuda de cara a mejorar la certificación de las causas de muerte. La discordancia puede surgir a lo largo del proceso que ha de seguirse en toda certificación de las causas de una defunción. Este proceso podemos estructurarlo en las siguientes fases, que ordenadas en el tiempo serían:

1.º adquisición de los datos clínicos. 2.º interpretación y elaboración de los diagnósticos. 3.º cumplimentación de los diagnósticos en el certificado de defunción. 4.º selección de la causa básica de defunción.

Hemos de pensar que las discrepancias difícilmente han surgido en la primera fase de adquisición de datos, pues ambos médicos disponían de idéntica información clínica, independientemente de que ésta fuera la deseable. Es en la fase de interpretación y elaboración de los diagnósticos donde creemos que radica la principal fuente de desacuerdo entre los dos médicos certificadores, pues en ella intervienen, además de las características de la información clínica dispo-

nible, cuya cantidad y precisión puede influir decisivamente en la interpretación que de los datos se haga, aquellos factores ligados a la formación y experiencia práctica de cada uno de ellos, que inevitablemente han de influir sobre la elaboración de los diagnósticos.

En este sentido, hay que destacar que el médico 2 es más exigente en cuanto a la valoración que hace de la calidad de los datos clínicos disponibles, apareciendo una importante discordancia entre ellos (Kappa 0.48), posiblemente debida a que al ser especialista en cardiología exige mayor precisión de los datos a la hora de catalogar una defunción como causada por alguna de las enfermedades circulatorias, siendo este grupo de causas, como vimos, donde se concentra el mayor número de defunciones en desacuerdo.

Las siguientes fases, cumplimentación y selección, estrechamente relacionadas entre sí no parecen que puedan explicarnos la discordancia encontrada. Ambos certificadores estaban bien instruidos en las normas a seguir para la cumplimentación correcta de las causas de la muerte, lo que facilita a su vez la selección de la causa básica por parte del codificador, que en nuestro estudio siempre fué, además, el mismo: por lo que tampoco surge discordancia que se deba a la variabilidad entre observadores⁶.

No obstante, los problemas derivados de la cumplimentación explican, como han señalado algunos autores⁷, gran parte de las discordancias que

aparecen en condiciones normales de certificación de causas de muerte y que nosotros creemos haber obviado.

Podemos concluir entonces, en primer lugar, que el diagnóstico estándar, tal como proponemos, puede ser el mejor de los posibles para validar las causas de muerte certificadas; en segundo lugar, que de darse unas condiciones similares a las existentes en nuestro estudio: (I) acceso a toda la información clínica existente por parte del médico que va a certificar la muerte, (II) entrenamiento en la correcta certificación de la causa de la muerte y (III) escasa variabilidad entre los codificadores que seleccionan la causa básica de muerte, las estadísticas de mortalidad serían mucho más fiables de lo que lo son en la actualidad.

Bibliografía

1. Fletcher RH., Fletcher SW., Wagner EH. *Clinical Epidemiology the essentials*. Baltimore: Ed. Williams and Wilkins, 1982.
2. Benavides FG. Fiabilidad de las estadísticas de mortalidad. *Monografías Sanitarias* 1986; 2 (serie A).
3. O.M.S. Manual de clasificaciones estadísticas internacional de enfermedades, traumatismos, y causas de defunción. *Publicaciones científicas*, 1978: 353.
4. Feinstein AR. *Clinical Epidemiology. The architecture of clinical research*. W.B. Saunders company, 1985.
5. Fleiss JL. *Statistical methods for rates and proportions*. John Wiley and sons, 1981.
6. Curb JD. et al. Nosological coding of cause of death. *American Journal of Epidemiology* 1983; 118(1):122-128
7. Domenech J., Segura A. El papel del médico en la calidad de las estadísticas de mortalidad. *Jano* 1984; octubre: 40-47.

