



Instituto Nacional de Salud Pública

**MARCO PROCEDIMENTAL PARA LA COMPARACIÓN DE LAS PRINCIPALES
CONDICIONES DE SALUD DE LOS DERECHOHABIENTES DEL ISSSTE, A
PARTIR DE LAS ENCUESTAS DE SALUD Y NUTRICIÓN, ENSADER 2007 Y
2012**

Francisco Alejandro López Carrillo
lopezcaralex@gmail.com
7773712787

Maestría en Salud Pública con área de concentración en
Bioestadística y Sistemas de Información en Salud

Generación 2011-2013

Director: Dr. José Moreno Montoya
Instituto Nacional de Salud Pública

Asesora: Dra. Lina Sofía Palacio Mejía
Instituto Nacional de Salud Pública

Cuernavaca Morelos, Agosto de 2013

Marco procedimental para la comparación de las principales condiciones de salud de los derechohabientes del ISSSTE, a partir de las encuestas de salud y nutrición, ENSADER 2007 y 2012

Resumen: Objetivo. Proponer un marco procedimental que permita comparar las enfermedades crónicas más relevantes en los derechohabientes del ISSSTE.

Material y métodos. Se emplearon las bases de datos de las ENSADER 2007 y 2012. El universo de estudio lo constituyeron los trabajadores activos, pensionados y sus respectivos beneficiarios mayores de 20 años (n= 11 983). Para describir el efecto de las variables explicativas sobre las de respuesta, se construyeron modelos de regresión logística. El análisis del efecto de colapsar categorías implicó categorizar la variable edad a partir de 7 categorías. Posteriormente, esta se colapsó gradualmente en 2 direcciones hasta dicotomizarla, los modelos resultantes se compararon por medio del criterio de información bayesiano, el log verosimilitud y el valor bajo la curva de Característica Operativa del Receptor. **Resultados.** Tres condiciones de salud muestran incrementos significativos en la morbilidad entre años a) Diabetes de 82 casos a 168 por 1000 derechohabientes. b) Enfermedades cardíacas de 45 a 60 por cada 1000, y c) cáncer de 1 a 2 por 1000 derechohabientes. Colapsar la variable edad representó variación en la magnitud de los parámetros estimados relacionada con la dirección en la que se realizó dicho colapso. **Conclusión.** Agrupar las categorías para facilitar el análisis estadístico puede modificar los resultados si se realiza sin previo análisis. La magnitud de los coeficientes de un modelo de regresión puede deberse a la real magnitud de los efectos y a los puntos de corte establecidos por el investigador.

Palabras Claves: colapso de categorías; análisis estadísticos; encuestas; métodos epidemiológicos;

Introducción

La literatura científica publicada en América Latina durante las dos últimas décadas revela que más de un 90% de la evidencia generada en salud pública, proviene de encuestas con diseños de corte transversal, ¹ por lo que, como herramienta de investigación, estos abordajes poblacionales constituyen una importante fuente de información epidemiológica en tanto que desempeñan la función crítica de medir la salud de una población, especialmente en los países en desarrollo. ^{2, 3} Entre los aspectos abordados por estas investigaciones se cuentan las evaluaciones de intervenciones para la prevención de enfermedades, de la atención curativa y en general de la calidad percibida por parte de los usuarios sobre los servicios recibidos, además, el reconocimiento de aspectos para el fortalecimiento del sistema sanitario e incluso temas de vigilancia epidemiológica. ² Así mismo, proveen información directa e indirecta sobre exposiciones o riesgos específicos de enfermedades de interés. ⁴

En México, este tipo de investigaciones han aportado significativamente para el reconocimiento del estado de salud de la población desde hace más de 25 años ⁵ y en particular con las Encuestas Nacionales de Salud (ENSA) 1987, 1994 y 2000; las Encuestas Nacionales de Nutrición (ENN) 1988 y 1999; y las Encuestas Nacionales de Salud y Nutrición (ENSANUT) 2006 y 2012, cuyos resultados constituyen uno de los principales sustentos para la consolidación del perfil epidemiológico nacional.

El Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado (ISSSTE), cuya población amparada a 2012 supera 12 millones de beneficiarios, ⁶ y la cual se caracteriza por poseer una estructura por edad más envejecida que el promedio nacional. ^{7 8} y con un aumento en la expectativa de vida en los seres humanos, la cual se asocia con un aumento rápido en el número total de adultos de mediana y avanzada edad, así como, con el incremento en la cifra de muertes debidas a enfermedades no transmisibles. ⁹ La Organización Mundial de la Salud (OMS) señala que a nivel global las enfermedades cardiovasculares son la principal causa de muerte por enfermedades no transmisibles con el 48%, seguidas del cáncer que ocasiona el 21%. El 3,5% de las defunciones debidas a

enfermedades no transmisibles son causadas por la diabetes, se prevé que el número anual de defunciones debidas a enfermedades cardiovasculares se incrementará de 17 millones en 2008 a 25 millones para el año 2030, y que las debidas al cáncer ascenderán de 7,6 millones a 13 millones en el mismo periodo de tiempo ⁹. En México las enfermedades transmisibles también son un problema de salud que se ha venido puntualizando durante las últimas décadas, ya que la enfermedad coronaria y la diabetes son las dos causas principales de muerte en nuestro país, seguidas del derrame cerebral.⁽¹⁰⁾ Institucionalmente dichos cambios en las frecuencias de enfermedad son también previsibles, por cuanto las poblaciones derechohabientes de cada sistema experimentan similares cambios en su distribución etaria, de esta misma manera, la población derechohabiente del ISSSTE atraviesa una situación semejante en tanto que mas de la mitad de su población, el 58% de sus afiliados esta constituido por personas de 45 años o más ⁷.

Con el propósito de contar con información específica y confiable sobre los aspectos de salud y nutrición de mayor relevancia para su población adscrita, el ISSSTE decide implementar la primera Encuesta Nacional de Salud y Nutrición del Derechohabiente del ISSSTE 2007 (ENSADER 2007) ⁷ esta encuesta se renueva para el año 2012 de forma anidada en la Encuesta Nacional de Salud y Nutrición, ENSANUT 2012 (ENSADER 2012), ⁸ sin embargo hasta el momento, no se ha realizado ningún análisis comparativo entre ellas y en consecuencia queda pendiente la tarea de valorar el cambio en las cifras de estas enfermedades, por lo que el contenido de este estudio pretende ser un primer paso en la comparación del estado de salud de los derechohabientes adultos del ISSSTE.

La comparación de estas encuestas, enfrenta un desafío en tanto que supone la completa comparabilidad de los resultados obtenidos entre ellas. Dicha posibilidad depende en gran medida de las dinámicas demográficas, la periodicidad de realización de los estudios, la forma como se recolecta la información, se miden y categorizan las variables, ¹² ¹³ entre otras.

Sobre la categorización de variables, diversos autores han señalado que en los análisis estadísticos de encuestas de salud, pueden presentarse efectos importantes sobre la precisión en la estimación de los factores asociados a las condiciones de salud de interés debido al colapso de categorías de las variables consideradas.^{14 15 16 17 18}

Por lo anterior, este trabajo tiene como propósito proponer un esquema de procedimiento que permita valorar los efectos derivados del colapso de categorías en las variables asociadas con las principales enfermedades crónicas en la población derechohabiente del ISSSTE mayores de 20 años.

Materiales y métodos

Se emplearon las bases de datos de las encuestas ENSADER 2007 y 2012. Ambas tuvieron como población de referencia los hogares de los afiliados con al menos un individuo derechohabiente del ISSSTE a nivel nacional. El universo de estudio se constituyó a partir de la totalidad de trabajadores activos, pensionados y sus respectivos beneficiarios (dependientes económicos amparados: cónyuge, hijos, ascendientes), de todos los grupos etarios y pertenecientes a las 35 delegaciones / entidades federativas del ISSSTE. Para el año 2007 el estudio implicó la entrevista de 11 013 personas. Por su parte, la muestra efectiva de la ENSADER 2012 fue de 5246 hogares, en los que se aplicaron 11 428 cuestionarios individuales. El marco muestral de estas investigaciones se fundamentó en la información de distribución poblacional desagregada por Áreas Geoestadísticas Básicas (AGEBs equivalentes a sectores censales urbanos), lo que fue complementado con un listado de dependencias con derechohabientes del ISSSTE y el número de derechohabientes reportado para cada dependencia. Se implementó en ambos casos un esquema de selección probabilístico, multietápico con selección proporcional al tamaño de los agebs.^{7 8}

Variables de comparación

Las comparaciones se establecieron mediante la construcción de modelos de

regresión. Se usaron como variables de respuesta aquellas preguntas en las encuestas que ofrecieran información sobre auto-reportes de diagnósticos de patologías, como son: enfermedades cardíacas (I00-79), diabetes Mellitus (E10-14), tumores malignos (C00-97) e hipertensión (D00-09) según códigos de la décima revisión de la clasificación internacional de enfermedades CIE 10.¹⁹ En la conformación de los modelos, dichas variables se categorizaron de manera dicotómica (0 valor de referencia para sujetos sanos). Para el caso de enfermedades cardíacas la pregunta utilizada en las dos encuestas fue: *¿le ha dicho el médico que usted tiene o tuvo ... a) infarto b) angina de pecho c) insuficiencia cardíaca y d) otra enfermedad del corazón.* Para determinar la condición de diabetes la pregunta a utilizar en las dos encuestas fue: *¿algún médico le ha dicho que tiene diabetes o el azúcar alto en la sangre?* Para considerar el caso de cáncer se utilizó la pregunta: *¿Alguna vez le ha dicho un médico que tiene o tuvo cáncer o un tumor maligno de cualquier tipo?* Para el caso de hipertensión la pregunta utilizada fue: *¿Algún médico le ha dicho que tiene la presión alta?* Para las variables independientes se consideraron las variables que pudieran estar asociadas según la evidencia reciente; con estilos de vida como consumo de tabaco²⁰ y consumo de alcohol,²¹ así como condiciones de salud asociadas a las enfermedades de interés, entre estas: dislipidemias,²² hipertensión y obesidad.^{23,24} Se incluyeron además las características de sexo y edad en años cumplidos.

Análisis estadístico

Acorde con la escala de medición de las variables, se calcularon medias, medianas y desviaciones estándar para las variables cuantitativas, mientras que para las variables categóricas se obtuvieron frecuencias de los valores de sexo y grupos de edad decenales desde los 20 hasta los 79 y mayores de 80. Para comparar las tasas de morbilidad entre un periodo y otro se realizó un ajuste de tasas usando como población estándar la población derechohabiente del ISSSTE 2012.²⁵

Para describir el efecto de las variables explicativas sobre las de respuesta, se construyeron modelos de regresión logística, primero seleccionando las variables independientes, las cuales fueron analizadas inicialmente en un modelo bivariado, para ser finalmente incluidas en el modelo definitivo según resultados obtenidos en la prueba de Wald ($P < 0.05$).^{26 27}

El análisis del efecto de colapsar deliberadamente las categorías en los modelos implicó inicialmente categorizar a la variable edad por cuanto esta fue incluida en las regresiones a partir de 7 categorías: 20-29, 30-39, 40-49, 50-59, 60-69, 70-79 y 80+. Posteriormente, esta se colapsó gradualmente en 2 direcciones (de la primera categoría hacia la última y viceversa) hasta dicotomizar la variable, formando de esta manera 11 grupos diferentes para la variable edad. Los modelos resultantes a partir de la incorporación de esta variable según nivel de categorización, se compararon por medio del Criterio de Información Bayesiano (BIC)²⁸ y la prueba de log verosimilitud;^{29 30} se utilizó también como criterio en la comparación el valor bajo la curva de Característica Operativa del Receptor (ROC).³¹ Los análisis se implementaron en el paquete estadístico STATA en su versión 11.³²

Resultados

Se analizó una muestra de 11 983 registros de personas mayores de 20 años, 4582 provenientes de la ENSADER 2007 y 7401 de la ENSADER 2012, que tras aplicar los factores de ponderación las cifras anteriores representan 3 124 118 asegurados para el 2007 y 7 933 259 para el 2012. Un total de 3061 mujeres (1 610 018 valor ponderado) fueron entrevistadas en el año 2007 y 4274 (4 321 580 valor ponderado) en el 2012. La distribución por grupo de edad y sexo se presentan en el cuadro 1.

Cuadro 1 Distribución de la muestra por grupo de edad y sexo. Valores ponderados

	2007						2012					
	Femenino		Masculino		Total		Femenino		Masculino		Total	
	F	%	F	%	F	%	F	%	F	%	F	%
20-29	116 354	7	109 272	7	225 626	7	908 778	21	932 372	26	1 841 150	23
30-39	419 982	26	354 855	23	774 837	25	840 775	19	608 988	17	1 449 763	18
40-49	490 790	30	343 848	23	834 638	27	821 435	19	757 537	21	1 578 972	20
50-59	339 133	21	367 078	24	706 211	23	839 942	19	623 297	17	1 463 239	18
60-69	168 202	10	229 723	15	397 925	13	450 076	10	363 362	10	813 438	10
70-79	57 454	4	82 319	5	139 773	4	314 225	7	225 427	6	539 652	7
80+	18 103	1	27 005	2	45 108	1	146 349	3	100 696	3	247 045	3
Total	1 610 018	100	1 514 100	100	3 124 118	100	4 321 580	100	3 611 679	100	7 933 259	100

Fuente ENSADER 2007 y ENSADER 2012

Las principales causas de morbilidad en los derechohabientes adultos del ISSSTE, exploradas en las encuestas ENSADER 2007 y 2012 fueron la hipertensión (19.0%), diabetes (11.6%), enfermedades del corazón (6.0%) y cáncer (2.2%) para 2012 (cuadro 2).

Cuadro 2 frecuencias de las enfermedades crónicas mas relevantes de los derechohabientes del ISSSTE por año de encuesta y sexo

	2007						2012					
	Hombres		Mujeres		Total		Hombres		Mujeres		Total	
	Casos	%	Casos	%	Casos	%	Casos	%	Casos	%	Casos	%
Hipertensión	322 743	21.3	304 819	18.9	627 562	20.0	593 693	16.4	921 498	21.3	151 5191	19.0
Diabetes	160 704	10.6	108 747	6.7	269 451	8.6	430 930	11.9	497 147	11.5	928 077	11.6
Cáncer	7 601	0.5	38 727	2.4	46 328	1.4	51 752	1.4	129 688	3.0	181 440	2.2
Corazón	68 439	4.5	62 238	3.8	130 677	4.1	208 530	5.7	273 059	6.3	481 589	6.0

Fuente: ENSADER 2007 Y ENSADER 2012

Población total hombres 2007 1 514 100, Población total mujeres 2007 1 610 018

Población total hombres 2012 3 611 679, Población total mujeres 2012 4 321 580

Estos resultados estandarizados para la comparación de las dos encuestas, de donde se pudo observar que la hipertensión arterial es la única condición que no ha aumentado significativamente (5%) desde el 2007 al 2012, de 190 a 200 casos por cada 1000 derechohabientes ($p=0.223$); las otras tres condiciones comparadas entre años muestran cambios estadísticamente significativos (con un valor $p < 0.05$) de la siguiente manera: a) Diabetes aumento de 82 casos a 168 por cada 1000 derechohabientes, significando un aumento del 104% ($p=0.000$). b)

Enfermedades del corazón aumento el 33% de 45 a 60 casos por cada 1000 derechohabientes, ($p=0.001$), y c) la tasa de morbilidad de cáncer aumento de 1 a 2 por cada 1000 derechohabientes, el 100% ($p=0.000$).

El efecto observado de las factores asociados a las condiciones de salud como lo es la variable edad, al colapsarla según diferentes categorizaciones representó una importante variación en la magnitud de los parámetros estimados y se relacionada con la dirección en la que se realizó dicho colapso, es decir, cuando la variable edad se colapsa de la última categoría a la primera los coeficientes se mantienen, en cambio, cuando la variable edad es colapsada de la primera categoría a la última, la magnitud de los coeficientes asociados con los parámetros estimados en el modelo disminuyen en la mayoría de los casos (figura 1).

Figura1 Descripción de los modelos

Modelos		
Variables dependientes	Variables independientes	Categorías en la variable edad
Diabetes	-Edad	M1 20-29, 30-39, 40-49,50-59,60-69,70-79,80+
	-Hipertensión	M2 20-39, 40-49,50-59,60-69,70-79,80+
	-Obesidad	M3 20-49,50-59,60-69,70-79,80+
	-Dislipidemias	M4 20-59,60-69,70-79,80+
Enfermedades del corazón	-Sexo	M5 20-69,70-79,80+
	-Antecedente padre	M6 20-79,80+
	-Edad	M7 20-29, 30-39, 40-49,50-59,60-69,70+
	-Hipertensión	M8 20-29, 30-39, 40-49,50-59,60+
Cáncer	-obesidad	M9 20-29, 30-39, 40-49,50+
	-Antecedentes padri	M10 20-29, 30-39, 40+
	-Dislipidemias	M11 20-29, 30+
	-Edad	
Hipertensión	-Obesidad	
	-Dislipidemias	
	-Diabetes	

Fuente: Marco procedimental para la comparación de las principales condiciones de salud de los derechohabientes del ISSSTE, a partir de las encuestas de salud y nutrición, ENSADER 2007 y 2012. Agosto 2013.

De esta manera podemos observar que en el análisis realizado para cáncer en 2012, los coeficientes disminuyen con el primer colapso (Modelo M2) en el grupo

de 80 y más años de 12.2 a 5.5, en el grupo de 70 a 79 años de 18.0 a 8.1, en el de 60-69 años, de 13.2 a 6.0, en 50-59 años, de 12.5 a 5.6 y en el de 40-49 años de 6.5 a 2.9, reduciendo un 55% el valor del coeficiente en cada una de las edades. En cuanto al cáncer en 2007, los coeficientes reducen en M2 el 30% en cada una de las categorías. En hipertensión en 2012 el M2 reduce el 36% en cada una de las categorías. En hipertensión 2007 es el caso en donde la reducción del odds al reagrupar las primeras categorías 20-39 es menos notorio cambiando 11%, sin embargo, cuando el colapso se realizó en sentido contrario no ocurrieron cambios significativos en los coeficientes de estas condiciones de salud. A modo de ejemplo se observan en el cuadro 3, la forma en que los odds ratios de los modelos de diabetes 2012 van cambiando colapso por colapso.

Cuadro 3 Descripción de los odds ratio a través de los colapsos en los modelos de diabetes 2012 (intervalos de confianza al 95%)

Grupo Edad	M1	M2	M3	M4	M5	M6
30-39	3.8 (1.7-8.7)					
40-49	14.4 (6.7-30.9)	5.4(3.8-7.7)				
50-59	27.3 (12.8-58.2)	10.3(7.4-14.4)	3.5(2.9-4.3)			
60-69	44.5 (20.8-95.4)	16.9(12.0-23.9)	5.8(4.7-7.2)	3.3(2.8-4.0)		
70-79	44.4(20.5-96.2)	16.9(11.7-24.4)	5.9(4.6-7.5)	3.4(2.7-4.3)	3.87(1.7-8.7)	
80+	41.0(18.3-91.6)	15.6(10.2-24.1)	5.5(3.9-7.7)	3.2(2.3-4.5)	26.9(12.7-57.0)	20.9(9.9-44.2)
	M1	M7	M8	M9	M10	M11
30-39	3.8(1.6-8.6)	3.8(1.6-8.6)	3.8(1.6-8.6)	3.8(1.71-8.7)	3.8(1.7-8.7)	20.9(9.9-44.2)
40-49	14.4(6.7-30.9)	14.4(6.7-30.8)	14.4(6.7-30.8)	14.5(6.8-31.2)	26.9(12.7-57.0)	
50-59	27.2(12.7-58.2)	27.2(12.7-58.2)	27.2(12.7-58.1)	35.2(16.6-74.6)		
60-69	44.5(20.7-95.4)	44.5(20.7-95.3)	43.9(20.6-93.5)			
70-79	44.4(20.5-96.2)	43.3(20.1-93.2)				
80+	41.0(18.3-91.6)					

Fuente: ENSADER 2012

Las reducciones en el número de categorías de la variable edad, arrojaron cambios en el log-verosimilitud de los modelos cuando fueron comparados con el modelo sin colapsar (M1), sin embargo, de la misma forma como ocurre con los parámetros estimados, estos cambios variaron según la dirección del colapso de las categorías; en general, cuando el colapso se realiza de la primera categoría a la última según orden de edad, la diferencia con el modelo sin colapsar se observa

estadísticamente significativa con menos ejercicios que cuando se realizan en dirección contraria. Así, en 6 ocasiones el log-verosimilitud del modelo reagrupado no es significativamente diferente al log-verosimilitud del modelo original usando el orden natural de la variable, mientras que en el análisis de la categorización inversa se observan 16 modelos colapsados que no son significativamente diferentes al modelo sin colapsar.

Desde la perspectiva de los cambios ocurridos en el valor BIC, se observó que la intensidad del efecto del colapso también se relaciona con la dirección en que estos se realizaron, haciéndolo más notable cuando las categorías se colapsaron de la más baja a la más alta, en comparación de cuando el colapso de categorías se realizó en la dirección contraria. La comparación a partir de este criterio se realizó tomando en cuenta la siguiente clasificación: ³³ Evidencia débil de un modelo mejor ajustado si la diferencia es de 0-2 unidades en el valor BIC; evidencia buena con una diferencia de 3-6; Evidencia fuerte si la diferencia en el valor BIC es de 7-10; y muy fuerte si la diferencia es > 10.

En relación al valor bajo la curva ROC como criterio de comparación, al igual que en los casos anteriores, su valor varía acorde a la dirección en que se colapsaron las categorías de la variable edad; así, las diferencias fueron mayores cuando las categorías se colapsaron de la primera en orden cronológico hacia la última, en comparación a cuando el colapso se realizó en el sentido contrario. Los detalles de esta comparación se presentan en el cuadro 4.

Cuadro 4 Efecto del colapso de la variable edad utilizada como variable independiente en los modelos de regresión logística

Variable de respuesta	Modelos	2007		2012	
		Diferencia BIC	Diferencia ROC	Diferencia BIC	Diferencia ROC
Diabetes	M2	5,39	0.08	-3,85	0
	M3	-12,81	0.00	-114,61	-0.01
	M4	-54,00	0.00	-272,81	-0.03
	M5	-128,66	-0.01	-416,36	-0.06
	M6	-154,86	-0.04	-473,61	-0.08
	M9	-13,93	0.00	-2,32	-0.00

Corazón	M10	-66,80	-0.03	-93,61	-0.02
	M11	-144,88	-0.05	-299,51	-0.06
	M2	5,14	0.00	2,97	-0.00
	M3*	13,31	-0.01	10,78	-0.00
	M4	14,40	-0.01	3,87	-0.01
	M5	-8,54	-0.04	-2,51	-0.01
	M6	-19,24	-0.05	-42,76	-0.03
	M8*	14,29	0.00	5,52	0.00
	M9	3,05	-0.01	-1,35	0.00
	M10	-14,54	-0.03	-24,89	-0.02
	M11	-26,39	-0.06	-53,32	-0.04
Cáncer	M2*	11,78	-0.00	3,59	-0.00
	M3*	18,32	-0.01	-3,31	-0.02
	M4	16,73	-0.06	-31,32	-0.06
	M5	12,02	-0.10	-36,33	-0.03
	M6	20,04	-0.10	-46,57	-0.12
	M10	24,88	-0.05	15,54	-0.03
	M11	22,88	-0.09	-4,65	-0.07
	M2*	8,03	0.00	-4,34	0.00
	M3	1,14	-0.00	-61,28	0.00
	M4	-67,12	-0.02	-212,70	-0.02
	M5	-208,48	-0.05	-473,53	-0.05
Hipertensión	M6	-275,84	-0.07	-689,48	-0.08
	M8*	12,92	-0.00	-4,41	0.00
	M9	-43,45	-0.01	-136,86	-0.01
	M10	-165,30	-0.04	-363,49	-0.04
	M11	-263,59	-0.07	-591,49	-0.08

* Sin diferencias significativas para la prueba de log-verosimilitudes en 2007

Fuente: ENSADER 2007 Y ENSADER 2012

Discusión

En este trabajo se compararon cuatro condiciones de salud en los derechohabientes del ISSSTE para los años 2007 y 2012 mediante los datos provenientes de las encuestas ENSADER correspondientes a estos años. En ambos casos se usó la información por auto-reporte relacionada con las principales enfermedades crónicas en la población, específicamente diabetes con una morbilidad de 168 por cada mil derechohabientes mayores de 20 años en 2012, enfermedad cardíaca 60 por cada mil, cáncer 2 por cada mil e hipertensión

con 20 casos por cada mil, todas dentro de las 10 primeras causas de mortalidad en los adultos del país.³⁴ Análisis preliminares sobre la frecuencia de estas enfermedades mostraron incrementos significativos en todas de la siguiente manera: diabetes en 104%, cáncer 100% y enfermedades cardíacas 33%. Hipertensión fue la única condición que no aumentó significativamente con el 5%.

El estudio más detallado de este tipo de fenómenos, no obstante, implica el involucramiento de más variables, situación por la cual es conveniente analizar el efecto que tiene la forma de medición de estas en la construcción de modelos de regresión. Por lo anterior, este trabajo buscó valorar el efecto que tiene la categorización arbitraria de variables continuas explicativas en modelos de regresión logística sobre las patologías antes mencionadas, en particular, la edad, tomando en cuenta su relación simultánea con todas las enfermedades crónicas analizadas.^{35 36}

Todos los análisis se desarrollaron agrupando arbitrariamente la variable edad medida en años cumplidos. Los resultados indican que los coeficientes de los modelos de regresión disminuyen significativamente en la mayoría de los casos al aumentar la edad de referencia, es decir, cuando el colapso se efectúa en dirección ascendente. Este resultado, acorde con lo verificado en la literatura previa, confirma que las personas de edad más avanzada observan indicadores de riesgo mayor; cuando se agrupan los individuos más jóvenes respecto a las siguientes categorías de edad, se logra una reducción en la diferencia de edades entre la categoría de referencia y las categorías a comparar, obteniendo como resultado una disminución del indicador de riesgo frente a la enfermedad analizada³⁷. En ambos casos, la magnitud de los coeficientes asociados al efecto de la variable puede reflejar al mismo tiempo, tanto la fuerza real de la asociación como la capacidad discriminativa del modelo en función del tamaño de muestra logrado en cada grupo, mejorando la eficiencia estadística del análisis.³⁸

Los modelos colapsados desde la primera categoría a la última según el orden natural de la variable, sugieren por tanto mejores niveles de ajuste, en particular

desde la perspectiva del criterio de información Bayesiano ³³ y el número de parámetros necesarios a estimar. En relación a la capacidad de discriminación de dichos modelos, esta no se ve afectada en tanto que los valores del área bajo la curva ROC ³⁹ no reflejan mayores diferencias.

Como es natural, los resultados de este estudio deben ser tomados en cuenta con reconocimiento de las limitaciones inherentes al tipo de estudio. De esta forma, es preciso reconocer que por tratarse de información secundaria, los resultados aquí presentados dependen de la calidad de los datos de cada encuesta. Así mismo, la información utilizada, basada en auto-reportes, limita la capacidad inferencial de los modelos. Finalmente, debido al uso de un diseño epidemiológico de corte transversal, el estudio de relaciones causales en la construcción de los modelos está descartado.

Dónde colocar los puntos de corte en las variables cuando estas tienen varias categorías requiere de atención y se debe evitar el uso de numerosos puntos de corte extremas. ¹⁸ Sin embargo, el atractivo de la agrupación de categorías para facilitar la interpretación y el análisis estadístico puede modificar de manera importante los resultados si se realiza sin previo análisis. ¹⁷ Si se ha decidido colapsar, este procedimiento debe hacerse procurando un número conveniente de muestra en cada una de las categorías y solo si el resultado es un modelo más eficiente que el modelo sin colapsar ³⁸ y en este no se afecta la significancia estadística de los parámetros ofrecidos por el modelo. En este trabajo encontramos positivo buscar uniformidad en los resultados de las categorías de riesgo, mediante el colapso de categorías de la variable edad, considerando al valor BIC como uno de los principales criterios de comparación, profundizando sobre la diferencia estadística de verosimilitudes y finalmente utilizando el valor debajo de la curva ROC como criterio de comparación en los casos donde encontramos un modelo colapsado con mejor BIC y sin diferencia significativa de verosimilitud con el modelo original.

Consideramos justificable el colapso de categorías, únicamente en los casos en los que detrás de un análisis de los efectos ocasionados por dicha práctica y

utilizando los criterios antes mencionados, los resultados fueran positivos para el modelo colapsado.

Con base en lo anterior, este estudio pretendió verificar el hecho de que la magnitud de los coeficientes de un modelo de regresión puede deberse de manera simultánea a la real magnitud de los efectos y a los puntos de corte establecidos por el investigador. Es por esto que la manipulación del número de categorías de una variable específica puede constituir un desafío importante en tanto que las conclusiones derivadas pueden verse afectadas. Así, además de las consideraciones de significancia estadística, la elección del número correcto de categorías debe permitir la correcta verificación de las hipótesis teóricas subyacentes al fenómeno estudiado y en consecuencia, soportarse principalmente en el conocimiento previo del fenómeno.

Por lo anterior, para la correcta consolidación de un marco procedimental de análisis en relación al número correcto de agrupaciones en una variable independiente, es necesario considerar al mismo tiempo el efecto específico de interés y la forma como este puede representarse adecuadamente a través de la clasificación de la población en grupos específicos conformados por la(s) variable(s) a incluir en los modelos. Esta situación sugiere entonces que la correcta interpretación derivada del ejercicio estadístico de regresión depende no solo de la calidad numérica de los modelos, sino de su adecuada formulación epidemiológica.

Bibliografía

1. Peruga , León , Child , Cruz , Hernández , Arredondo , et al. Análisis de la participación en encuestas de cinco países: relevancia para la investigación en salud pública. Revista Panamericana de Salud Pública. 2000 Abril; 7(4).
2. Organización Panamericana de la Salud. Métodos de Encuesta para Lugares o Circunstancias Dificiles. Rev. Panamericana Salud Publica. 2007 Junio; 21(6).
3. SEDLAC. Socio-Economic Database for Latin America and the Caribbean. [Online].; 2012 [cited 2012 noviembre 30. Available from: <http://sedlac.econo.unlp.edu.ar/esp/index.php>.
4. Hernandez B, Velasco H. Encuestas Transversales. Salud Pública de México. 2000 septiembre-octubre; 42(5).
5. Instituto Nacional de Salud Pública. insp.mx. [Online]. [cited 2013 Mayo 22. Available from: <http://www.insp.mx/encuestoteca.html>.

6. Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado. Sitio web del Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado. [Online]. [cited 2013 Mayo 22. Available from: <http://www.issste.gob.mx/issste/anuarios/>.
8. Instituto Nacional de Salud Pública. Encuesta de Salud y Nutrición de los Derechohabientes del ISSSTE (ENSADER 2012) anidada en la Encuesta de Salud y Nutrición (ENSANUT 2012). ENSADER 2012. México; 2012.
7. Instituto Nacional de Salud Pública. Encuesta Nacional de Salud y Nutrición de los Derechohabientes del ISSTE 2007. Encuesta. Cuernavaca: INSP; 2007.
9. Organización Mundial de la Salud. Estadísticas sanitarias mundiales 2012. Suiza; 2012.
10. Villalpando S RR. Estatus de las enfermedades crónicas no transmisibles en México en base a la Encuesta Nacional de Salud y Nutrición 2006. Salud pública. 2010; 52(1).
11. Kuri P. La transición en salud y su impacto en la demanda de servicios. Gaceta Médica de México. 2011 diciembre; 147(6): p. 451-4.
12. Damián A. Los problemas de comparabilidad de la ENIGH y su efecto en la medición de la pobreza. Papeles de población. 2007;(51): p. 111-46.
13. Swart E. The prevalence of medical services use. How comparable are the results of large-scale population surveys in Germany? GMS Psycho-Social-Medicine. 2012 octubre; 9(10).
14. Kateri M, Iliopolus G. On collapsing categories in two-way contingency tables. AMS. 2000; 62(17).
15. Manor O, Matthews S, Power C. Dichotomous or categorical response? Analysing self-rated health and lifetime social class. International Journal of Epidemiology. 2000; 29: p. 149-57.
16. Lash T. Lash collapsing high-end categories of comorbidity may yield misleading results. Clinical Epidemiology. 2009; 1: p. 11-15.
18. Stromberg U. Collapsing Ordered Outcome Categories: A Note of Concern. American Journal of Epidemiology. 1996; 144(4): p. 421-4.
17. Ahern T, Bosco J, Silliman R, Ulcickas M, Campo T, Wei F, et al. Potential misinterpretations caused by collapsing upper categories of comorbidity indices: An illustration from a cohort of older breast cancer survivors. Clinical Epidemiology. 2009 Agosto; 1: p. 93-100.
19. CIE10. CIE10.org. [Online]. [cited 2013 Mayo 31. Available from: http://cie10.org/Cie10_Buscar_Consultar_En_Linea.php#PorCaps.
20. Samet M. Los riesgos del tabaquismo activo y pasivo. salud pública de México / vol.44, suplemento 1 de 2002. 2002; 44(1).
21. Medina-Mora ME, Natera G, Borges G, Cravioto , Fleiz C, Tapia-Conye. DEL SIGLO XX AL TERCER MILENIO. LAS ADICCIONES Y LA SALUD PÚBLICA: DROGAS, ALCOHOL Y SOCIEDAD. Salud Mental. 2001 Agosto; 24(4).
22. Soca. Dislipidemias. ACIMED 2009; 20(6): p. 265-273.
23. Córdova-Villalobos J, Barriguete-Meléndez A, Lara-Esqueda , Barquera , Rosas-Peralta , Hernández-Ávila , et al. Las enfermedades crónicas no transmisibles en México: sinopsis epidemiológica y prevención integral. Salud pública Méx . 2008 Octubre; 50(5).
24. Rojas-Martínez R, Aguilar-Salinas C, Jiménez-Corona A, Gómez-Pérez F, Barquera S, Lazcano-Ponce E. Prevalence of obesity and metabolic syndrome components in Mexican adults without type 2 diabetes or hypertension. Salud Publica Mex. 2012 Jan-Feb;54(1):7-12.. 2012 Jan-Feb; 54(1): p. 7-12.
25. PAHO. La Estandarización: Un Método Epidemiológico Clásico para la Comparación de Tasas. Boletín Epidemiológico / OPS. 2002; 23(3).
26. UCLA Institute for digital Research And Education. Stata Data Analysis Examples Logistic Regression. [Online]. [cited

- 2013 Junio 4. Available from: <http://www.ats.ucla.edu/stat/stata/dae/logit.htm>.
28. KENNETH P. BURNHAM , DAVID R. ANDERSON. Understanding AIC and BIC in Model Selection. *SOCIOLOGICAL METHODS & RESEARCH*. 2004 Noviembre; 33(2).
27. Park HA. An Introduction to Logistic Regression: From Basic Concepts to Interpretation with Particular Attention to Nursing Domain. *J Korean Acad Nurs* Vol.43 No.2 April 2013. 2013 Abril; 43(2): p. 154-164.
29. Williams R. University of Notre Dame Richard Williams Sociology Stata Highlights. [Online]. [cited 2013 Mayo 1. Available from: <http://www3.nd.edu/~rwilliam/stats2/l83.pdf>.
30. Statacorp. Stata: Release 12. Statistical Software. Station C, editor. Texas: StataCorp LP; 2011.
31. Fawcett T. ROC Graphs: Notes and Practical Considerations for Researchers. HP Laboratories. 2004 Marzo; 12(56).
32. StataCorp. [Online].; 2009.
33. Williams R. University of Notre Dame Richard Williams Sociology Stata Highlights. [Online]. [cited 2013 Mayo 5. Available from: <http://www3.nd.edu/~rwilliam/stats3/L05.pdf>.
34. DGIS. dgis.salud.gob.mx. [Online]. [cited 2013 abril 19. Available from: http://www.dgis.salud.gob.mx/descargas/pdf/ProNaSa_2007-2012.pdf.
35. Ordúñez García; P, Pérez Flores E, Hospedales J. High blood pressure care: beyond the clinical setting. *Rev Panam Salud Publica*. ; 28(4).
36. Gaziano , Bitton , Anand S, Abrahams-Gessel , Murphy. Growing Epidemic of Coronary Heart Disease in Low- and Middle-Income Countries. *Curr probl Cardiol*. 2010 Febrero; 35(2).
38. Marinelli , Winzer. agrupamiento de filas y columnas homogéneas en modelos de correspondencia. *Revista de Matemática: Teoría y Aplicaciones*. 2004; 11(1): p. 58-69.
37. Dominguez-Almendros S, Benitez-Parejo N, Gonzalez-Ramirez A. Logistic regression models. *Allergol Immunopathol*. 2011; 39(5).