

Fraudes a los sistemas de salud en Chile: un modelo para su detección

Francisco R. Mesa,^{1,2} Andrés Raineri,³ Sergio Maturana²
y Ana María Kaempffer⁴

Forma de citar

Mesa FR, Raineri A, Maturana S, Kaempffer AM. Fraudes a los sistemas de salud en Chile: un modelo para su detección. Rev Panam Salud Publica. 2009;25(1):56-61.

RESUMEN

Objetivos. Desarrollar un modelo para detectar casos de fraude planificado en Chile a partir de los datos contenidos en los formularios de licencia médica curativa (LMC) y establecer la contribución relativa de esos datos a su detección.

Métodos. Se aplicó un modelo de regresión logística binominal a partir de cuatro variables contenidas en el formulario legal de LMC exigido nacionalmente para justificar las ausencias al trabajo por motivos de enfermedad: el número de licencias médicas asignadas a una misma persona, el número de días de licencia médica otorgados por el médico tratante, el monto total a pagar por la enfermedad y una variable dicotómica que refleja si el diagnóstico es comprobable o no. Se analizaron 4 079 LMC presentadas el año 2003 a una institución privada de salud previsual, de las cuales 356 estaban ya clasificadas como fraudulentas por un panel de médicos expertos en fraude.

Resultados. El modelo logró identificar correctamente 99,71% de las licencias médicas fraudulentas y 99,86% de las no fraudulentas según el criterio del panel de expertos en fraude. Tres de las variables empleadas presentaron un poder predictivo independiente estadísticamente significativo. El valor predictivo positivo del modelo propuesto fue de 98,59%, mientras el valor predictivo negativo fue de 99,97%.

Conclusiones. El modelo logístico binomial desarrollado, basado en cuatro variables de uso universal en los formularios de LMC utilizados por todas las entidades aseguradoras de Chile, tanto públicas como privadas, permite discriminar de forma precisa y más rápidamente y con menor costo que los paneles de expertos las solicitudes fraudulentas de las no fraudulentas.

Palabras clave

Fraude, seguro de salud, técnicas de apoyo para la decisión, Chile.

En la literatura científica se distinguen diversos tipos de fraudes a los sistemas

de seguro de salud. Los llamados fraudes planificados son actos criminales del asegurado al tratar de obtener dinero u otra compensación por medio de declaraciones o evidencias físicas falsas (1, 2). En el fraude oportunista, el asegurado trata de magnificar sus daños con la finalidad de obtener un mayor pago por un evento (1), lo que en el caso de la salud correspondería a la sobreutilización de un beneficio. Estos fraudes se han documentado en todo el mundo, ya sea en países industrializados como los Estados Unidos de América (2-4), Aus-

tralia (5) y España (6, 7), o en países en desarrollo como Ghana (8).

Los fraudes a los sistemas de salud y sus crecientes costos preocupan a todos los actores involucrados en esta actividad, tanto a auditores y fiscalizadores (9-10) e investigadores académicos (11) como a los administradores de los sistemas de salud públicos (5, 7, 11) y privados (2-4). Algunos autores sostienen que en los Estados Unidos, el fraude planificado puede llegar a 5% del gasto total en salud (2). Un estudio sobre seguros médicos por lesiones en accidentes de tránsito encontró

¹ Universidad de Los Lagos, Departamento de Gobierno y Empresa, Puerto Montt, Chile. La correspondencia se debe dirigir a Francisco Mesa, Universidad de Los Lagos, Departamento de Gobierno y Empresa, Camino Chiquihue km 6, Puerto Montt, Chile. Correo electrónico: fmesa@ulagos.cl

² Pontificia Universidad Católica de Chile, Departamento de Ingeniería Industrial y Sistemas, Santiago, Chile.

³ Pontificia Universidad Católica de Chile, Escuela de Administración, Santiago, Chile.

⁴ Universidad de Chile, Escuela de Salud Pública, Santiago, Chile.

que 72% de las solicitudes presentadas se cancelaron inmediatamente, 27,7% se clasificaron como sospechosas y se alcanzó un acuerdo entre la aseguradora y los asegurados, y solo 0,3% de los casos terminaron con un veredicto de culpabilidad en los tribunales (4). En otra investigación publicada en 2002 se informó que de 17 274 casos sospechosos de fraude, en 3% los tribunales llegaron a un veredicto de culpabilidad y en 1% el juicio quedó pendiente, por lo que los casos reales de fraude podrían llegar a 4% del total de solicitudes de pagos por gastos de salud en esa población (2).

Por su parte, las estimaciones de la sobreutilización de los beneficios o fraudes oportunistas en los sistemas de salud varían según el país. Fuentes gubernamentales de los Estados Unidos han señalado que 10% del gasto total en salud contiene elementos de fraude (2, 12), en España la sobreutilización del seguro de enfermedad podría llegar a 14% (7) y en Chile las estimaciones ascienden a 27% (13).

El interés académico en los fraudes a los sistemas de salud es escaso y relativamente reciente —los primeros informes datan de la década de 1980 (2)— y para la detección y estimación de los casos de fraude se han utilizado pocas herramientas informáticas y estadísticas (2–5). Hasta el momento no se han dado a conocer explícitamente detalles de los modelos y métodos de detección empleados, ya que esa información podría ayudar a evadir la detección (12).

Métodos empleados para detectar fraudes

Entre las herramientas más utilizadas por los autores y los investigadores para detectar fraudes al sistema de salud se encuentran las redes neuronales artificiales (RNA) y el análisis de regresión multifactorial. Ambos métodos se basan en el aprendizaje supervisado de esos sistemas que se van ajustando con los datos de casos previamente clasificados como fraudulentos o no (14) y cada uno posee sus propias condiciones de uso, fortalezas y debilidades. En el caso de las RNA, su principal ventaja es su rapidez, pero no permiten identificar la contribución específica de cada variable independiente, es decir, el usuario no conoce el modelo o algoritmo que se utiliza y no recibe una explicación de por qué un caso se clasifica como fraudulento o no (14). Por su parte, la principal ventaja del análisis

de regresión multifactorial es que permite establecer la contribución relativa de cada una de las variables a la conclusión final de un posible fraude, pero es más lento ya que requiere una mayor comprensión del fraude efectuado y se debe realizar una correcta selección de las variables más significativas.

En Australia (5) se han utilizado RNA para clasificar las prácticas de atención de los médicos generales y detectar acciones inadecuadas —como realizar más servicios de los que se precisan o visitar a los pacientes más de lo necesario— que llevarían a descubrir casos de fraude. Aunque en ese modelo se utilizaron 28 variables, estas no se mencionaron por razones legales y profesionales, según los autores. En España, el empleo de las RNA en el control del gasto público por incapacidad temporal ha permitido un mejor control y eficiencia del gasto en salud y el aumento en la satisfacción percibida por los asegurados (7).

El análisis de regresión multifactorial se ha utilizado en un solo caso documentado de detección de fraude (4). En esa investigación se analizaron las solicitudes de reembolsos por lesiones físicas en accidentes de tránsito a una compañía de seguros de los Estados Unidos. Mediante el análisis de componentes principales se asignaron puntuaciones a un conjunto de variables de predicción obtenidas del documento de reembolso y luego un programa computacional clasificó las solicitudes de reembolsos en un grupo de fraude y otro de no fraude, según la probabilidad de fraude estimada por el modelo. Las variables de predicción más destacadas fueron: el tener un historial previo de reembolsos, no tener evidencia objetiva de la lesión y las características del salario. Los resultados del estudio indican que ese modelo tuvo una capacidad predictiva similar y fue menos costoso que la utilización de investigadores de seguros.

Una tercera metodología empleada para la detección de fraudes en salud es la llamada heurística, basada en el análisis de indicadores (3). Los métodos heurísticos combinan la información financiera con datos del proveedor y elementos de lógica médica y logística. Mediante uno de estos métodos, el de detección electrónica de fraude, se analizaron las solicitudes de reembolso de gastos médicos en una compañía de seguros de salud de los Estados Unidos y se compararon con datos estadísticos previos de fraude y el

comportamiento de los proveedores del sistema de salud que habían cometido fraudes. Para ello se clasificó el comportamiento de los proveedores de salud a partir de 27 indicadores agrupados en cinco categorías: aspectos financieros de la solicitud, lógica médica (normal o anormal), frecuencia del tratamiento, aspectos logísticos (lugar, tiempo y secuencia de actividades) e identificación del proveedor del seguro.

Según una revisión de cinco trabajos sobre la detección de fraudes (2), otro método empleado para obtener conclusiones generales acerca de los procedimientos utilizados por las empresas aseguradoras se basa en el análisis del flujo de las actividades comunes de esas empresas y la identificación de posibles variables de predicción para utilizar en una posterior modelación. Se encontró que algunos indicadores del reembolso de seguros permitían pronosticar fraudes y construir modelos que evitarían errores de pronóstico, sean estos falsos positivos o falsos negativos. En estos modelos se incorporaron variables subjetivas o procesadas, ya que aportan relaciones significativas asociadas con la capacidad de predicción de esos modelos. De esas variables, las más atractivas para aplicar en este estudio fueron tener un historial previo de reembolsos, no tener evidencia objetiva de la lesión y las características del salario (4).

Al desarrollar nuevos métodos para identificar casos potenciales de fraude se deben tomar en cuenta las restricciones reales con que operan los sistemas de salud, en particular la falta y asimetría de la información sobre los asegurados, su estado de salud y la entidad aseguradora (2).

Particularidades del sistema de salud chileno

Chile ha sido líder y pionero en la introducción de sistemas privados de salud y seguridad social. Su experiencia ha sido objeto de estudio y análisis en el ámbito académico, tanto por organismos internacionales como por gobiernos (15, 16). El sistema de salud chileno está compuesto por el sector público —que otorga prestaciones a 72,5% de la población, con un administrador financiero denominado Fondo Nacional de Salud (FONASA) y un sistema de prestadores públicos— y el sector de seguros de salud privado, formado por instituciones

de salud previsional, conocidas como “isapres” —utilizado por 16,8% de la población, principalmente personas de mayores ingresos—; el 10,7% restante corresponde a los miembros de las Fuerzas Armadas y a personas no cubiertas por ningún sistema (17).

El sistema de isapres, nacido en 1981, es el resultado de modificaciones legales que permitieron a empresas privadas proveer seguros de salud y normaron su funcionamiento (18). En la actualidad coexisten dos tipos de isapres: las abiertas y las cerradas. En las primeras (95,5% del total de cotizantes de las isapres) puede ingresar cualquier trabajador. En este segmento, ocho isapres compiten libremente por las primas de salud de los trabajadores. En las isapres cerradas (4,5% del total de cotizantes) solo pueden ingresar los trabajadores de las empresas que poseen la capacidad financiera para tener isapres de este tipo. En diciembre de 2006, el sistema de salud privado atendía a 1 286 165 trabajadores cotizantes, con 1 398 389 asegurados adicionales, en su mayoría cónyuges e hijos (19).

El sistema de salud chileno permite justificar ausencias laborales por tres causas: enfermedad del trabajador; embarazo, parto y puerperio de la madre trabajadora; y enfermedad grave de un hijo menor de un año. En esos casos, el trabajador obtiene una licencia con un subsidio equivalente a 100% de su salario o remuneración mientras dure la enfermedad o situación que originó la licencia. Este subsidio, que desde el punto de vista económico es un seguro, solo se puede obtener cuando un médico cirujano, un dentista o una matrona certifican mediante un documento de licencia médica curativa (LMC) que el trabajador está incapacitado para laborar debido a una de las situaciones anteriores (20, 21). Este documento tiene valor legal y contiene los datos del trabajador enfermo, además de la cantidad de días de reposo que necesita para recuperar su salud (21).

El hecho de que los sistemas de salud público y privado en Chile estén obligados por la ley a utilizar el Formulario Nacional de LMC convierte este documento en un instrumento válido para identificar y pronosticar posibles casos de fraudes relacionados con estas licencias.

El objetivo de la presente investigación fue desarrollar un modelo para detectar casos de fraude planificado en Chile a partir de los datos contenidos en los formularios de LMC y establecer la

contribución relativa de esos datos a su detección.

MATERIALES Y MÉTODOS

Se realizó un estudio predictivo en una compañía del tipo isapres que opera en todo Chile, a la que están afiliados más de 250 000 cotizantes (20% del total de los asegurados por las isapres en el país) con una cobertura total de más 550 000 beneficiarios si se toman en cuenta los asegurados adicionales (hijos, cónyuges y otros familiares incluidos en las pólizas) (19).

Para este estudio se analizaron los 4 082 formularios de LMC presentados a esa compañía entre el 1 de enero y el 31 de diciembre de 2003. De ellos, 356 estaban clasificados como fraudulentos mediante el proceso habitual de análisis llevado a cabo por esa compañía de seguro. Ese proceso se basa en el análisis de médicos expertos en fraudes, los contratos de afiliación, los formularios de LMC, los antecedentes de la empresa aseguradora y el cotizante en diversas bases de datos, la verificación del diagnóstico, los resultados de las visitas hechas a la empresa del cotizante para verificar la existencia de ambos y, finalmente, la confirmación o no de la existencia de la empresa a partir de los datos del Servicio de Impuestos Internos. En algunos casos se tomaron también en cuenta documentos judiciales y la información proporcionada por la Policía de Investigaciones de Chile. Muchos de los fraudes detectados se cometieron mediante empresas ficticias creadas con el fin de cobrar las LMC mientras en otros casos se trataba de personas que no estaban realmente empleadas.

Los datos contenidos en los formularios de LMC se digitaron en una base de datos y se utilizaron como variables independientes. Como dependiente se empleó la variable dicotómica de caso fraudulento o no fraudulento. La base de datos fue proporcionada por la compañía aseguradora y está disponible para los organismos reguladores del sistema de salud chileno (20).

Método y variables

Se aplicó el análisis de regresión logística binomial, —una variante de la regresión logística multifactorial utilizada cuando la variable dependiente es dicotómica como en este caso— a partir de cuatro variables seleccionadas del formulario de LMC que ya se han emple-

ado para detectar fraudes a sistemas de salud. Esas variables son las siguientes:

- días de licencia médica otorgados por el médico tratante; esta variable se empleó en un estudio anterior en una compañía de seguro del tipo isapres en Chile (22)
- monto a pagar por los días de enfermedad; esta variable se utilizó en un estudio de detección de fraudes realizado en los Estados Unidos bajo la definición de “características del salario” (4)
- clasificación del diagnóstico de la enfermedad por la que se solicita la LMC, según la posibilidad de comprobar ese diagnóstico mediante exámenes independientes. Médicos expertos en la detección de fraudes asignaron el valor 1 a los diagnósticos que pueden ser comprobados, como las fracturas, y el valor 0 a las enfermedades que se pueden simular y para las que no hay pruebas diagnósticas objetivas, como el lumbago o la depresión
- número de LMC otorgadas a una misma persona; esta variable es similar al historial previo de reembolsos del seguro de salud utilizado en otro estudio (4).

Con esas variables se procedió a construir el siguiente modelo de regresión logística binomial:

$$PFLMC = \beta_0 + (\beta_1 D) + (\beta_2 M) + (\beta_3 C) + (\beta_4 N)$$

donde:

- PFLMC: probable fraude en la LMC
 β_0 : constante de la regresión
 β_{1-4} : coeficientes beta específicos para cada una de las cuatro variables independientes
 D: días de licencia médica otorgados por el médico tratante
 M: monto a pagar en dinero por los días de enfermedad
 C: clasificación del diagnóstico
 N: número de licencias médicas que acumula la misma persona.

Se utilizó el programa estadístico SPSS v. 12 para estimar los valores de los coeficientes beta de la ecuación (ver valores numérico en el cuadro 1).

Para evaluar la bondad del ajuste del modelo se utilizó la prueba de Hosmer y Lemeshow (23). La significación estadística del impacto de cada variable dependiente sobre la probabilidad de que un

CUADRO 1. Peso y significación de cada variable del modelo binomial en la predicción de fraude en las licencias médicas clínicas (LMC) en la entidad estudiada

Variable	Coefficiente β^a	Error estándar	OR ^b	IC95% ^c
Categoría del diagnóstico	4,448	1,305	85,47	6,62–1 104,01
Total a pagar	$8,73 \times 10^{-6}$	$3,12 \times 10^{-6}$	1,00	1,00–1,00
Días de licencia	0,078	0,027	1,08	1,03–1,14
LMC por persona	2,306	0,345	10,04	5,10–19,73
Constante de la regresión, β_0	-10,883	1,516	NA ^d	NA

^a Coeficiente β : coeficientes calculados para cada variable independiente del modelo de predicción.

^b OR: razón de posibilidades (*odds ratio*).

^c IC95%: intervalo de confianza de 95%.

^d NA: no aplica.

CUADRO 2. Resultados de la valoración de las licencia médicas clínicas (LMC) según el modelo propuesto y el panel de médicos expertos en fraudes en la entidad estudiada

	Clasificación de las LMC según el panel de médicos expertos en fraudes			
	Fraude	No Fraude	Total	
Clasificación de las LMC según el modelo propuesto	Fraude	351 (A)	5 (B)	356 (A + B)
	No Fraude	1 (C)	3 722 (D)	3 723 (C + D)
	Total	352 (A + C)	3 727 (B + D)	4 079 (N)

Prevalencia: $[(A + C) / N] = 8,60\%$.

Sensibilidad: $[A / (A + C)] = 99,71\%$.

Especificidad: $[D / (B + D)] = 99,86\%$.

Valor predictivo positivo: $[A / (A + B)] = 98,59\%$.

Valor predictivo negativo: $[D / (C + D)] = 99,97\%$.

caso sea fraudulento se determinó mediante la razón de posibilidades (*odds ratios*, OR) con sus intervalos de confianza de 95% (IC95%). Se empleó un nivel de significación $P \leq 0,05$.

RESULTADOS

Para el análisis se utilizaron los datos de 4 079 formularios de LMC contenidos en la base de datos estudiada; se eliminaron 3 formularios (0,07% del total) por falta de información.

El modelo con las cuatro variables escogidas permitió identificar las LMC fraudulentas. Las variables categoría de diagnóstico y número de LMC por persona tuvieron una capacidad de predicción fuerte, mientras las variables días de licencia y suma total a pagar presentaron un poder de predicción significativo marginal (cuadro 1). Se comprobó la ausencia de multicolinealidad entre las variables independientes, ya que los valores de correlación encontrados entre ellas fueron muy bajos (entre 0,02 y 0,32).

La prueba de Hosmer y Lemeshow (23) demostró que el modelo presentado en

este estudio se ajusta adecuadamente a los datos (grados de libertad = 8; $\chi^2 = 1,24$; $P = 0,996$). De esta forma, los rasgos más característicos de estos fraudes fueron la existencia de múltiples peticiones de licencias médicas para una misma persona y los diagnósticos de difícil comprobación. Además, en la medida en que aumentó el número de días de licencia médica, fue mayor la probabilidad de que el caso correspondiera a un fraude (cuadro 1).

En el cuadro 2 se comparan los resultados del modelo utilizado con el dictamen de los médicos expertos, considerado hasta el momento como el mejor criterio de referencia. La mayoría de los casos fueron clasificados correctamente por el modelo y 8,6% de las LMC presentadas se identificaron como fraudulentas. La sensibilidad y especificidad del modelo fueron muy elevadas (99,71% y 99,86%, respectivamente).

El valor predictivo positivo (es decir, la proporción de LMC fraudulentas detectadas correctamente por el modelo) fue de 98,59%, mientras el valor predictivo negativo (es decir, la proporción de LMC no fraudulentas que el modelo consideró

sin fraude) ascendió a 99,97% (cuadro 2). De esta manera, se puede concluir que el modelo tuvo un desempeño satisfactorio con un alto porcentaje de aciertos.

DISCUSIÓN

En Chile ya se había desarrollado un modelo basado en RNA para detectar fraudes y la utilización excesiva de LMC en una compañía de seguro del tipo isapres. Esas redes se nutrieron de cuatro fuentes de información: los formularios de LMC y los datos del trabajador beneficiario, el empleador y el médico que prescribió la licencia. Ese modelo permitió calcular la probabilidad de fraude de cada licencia emitida con una disminución de 10% en los costos de detección de fraudes; además permitió reducir en seis meses el tiempo necesario para detectar los posibles casos de fraude (22).

El modelo desarrollado aquí para la detección de fraudes en las LMC a partir de los datos de una compañía de seguro chilena resultó un instrumento eficaz y además práctico, ya que se puede aplicar tanto en el sector público como privado, con transparencia en las variables utilizadas. Este es el primer modelo de este tipo basado en las cuatro variables seleccionadas y, dada la universalidad de ellas, podría aplicarse en otros países donde se disponga de datos similares.

Este modelo está basado en el análisis de regresión logística binomial (4, 14), un método de análisis multifactorial no utilizado con anterioridad para detectar fraudes a los sistemas de salud. Sus resultados son novedosos y robustos. Además, este método de análisis puede emplearse cuando los supuestos de otras técnicas estadísticas no se cumplen, como la linealidad entre las variables dependientes e independientes, la normalidad en la distribución de las variables (este supuesto no se cumple cuando al menos una variable es cualitativa) y homocedasticidad de las varianzas (23, 24).

Los elevados valores de sensibilidad y especificidad del modelo y su valor predictivo —tanto positivo como negativo— cercano al máximo posible demuestran la utilidad del modelo propuesto para detectar con un alto grado de certeza las LMC con indicios de fraude. Además, al utilizar solamente cuatro variables que mostraron un adecuado poder de predicción, este modelo tuvo un mayor poder explicativo de cada variable que el modelo basado en RNA.

Entre las ventajas del modelo desarrollado aquí en comparación con otros métodos está su alta eficacia en la identificación de los casos de LMC fraudulentas, su bajo costo de implementación, su aplicabilidad tanto en instituciones de salud públicas como privadas y la rapidez con que se puede realizar el análisis y hacer los pagos de las solicitudes no fraudulentas. La implementación de este método puede disuadir a potenciales interesados en cometer fraudes y elevar el grado de satisfacción de los médicos tratantes al no cuestionarse en exceso sus diagnósticos de licencias médicas e, indirectamente, su actuación profesional.

La compañía de seguro estudiada ha estado afectada por distintos tipos de fraude, el más relevante por sus costos está relacionado con los pagos por LMC fraudulentas. Si se toma en cuenta que el máximo tope a cancelar es de US\$ 2 629 mensuales, un solo fraude en LMC puede representar anualmente un monto de US\$ 31 548, cifra muy relevante si se compara con el resto de las prestaciones.

El intervalo de confianza para la variable monto a pagar indica que no existe una asociación significativa entre ella y la variable dependiente (cuadro 1), pues los montos a pagar varían ampliamente, tanto en las LMC fraudulentas como en

las no fraudulentas. No obstante, se decidió mantener esta variable en el modelo propuesto debido a que otros autores han sugerido su importancia en la detección de fraudes en salud (4). El factor de predicción más significativo en este estudio, la cantidad de licencias solicitadas por una misma persona, influye indirectamente en el monto a pagar debido a la acumulación o suma de muchas LMC independientes.

Al analizar estos resultados se debe tener en cuenta que los datos tomados para este estudio podrían no ser totalmente representativos de los fraudes cometidos en las LMC en otros años o instituciones. Al no contar con series de otros años o datos de otras instituciones, este punto no puede confirmarse. Por otra parte, podría haber otros tipos y modalidades de fraude que aún los paneles de experto no hayan detectado y que por lo tanto pueden haberse pasado por alto en este análisis.

A pesar de estas limitaciones, se puede afirmar que el modelo logístico binomial desarrollado, basado en cuatro variables de uso universal en los formularios de LMC utilizados por todas las entidades aseguradoras de Chile, tanto públicas como privadas, permite discriminar de forma precisa las solicitudes fraudulen-

tas de las no fraudulentas y lo hace más rápidamente y con menor costo que los paneles de expertos.

Se debe ampliar este estudio a otros años e instituciones con el fin de corroborar y ajustar los resultados obtenidos. Dada la escasez de recursos y la inflación en los costos de todos los sistemas de salud en el mundo, estudios como el presentado aquí cobran especial importancia para los sistemas nacionales de salud y deben complementarse con análisis comparativos a nivel internacional.

Agradecimientos. Los autores agradecen el apoyo de la compañía de seguro que nos facilitó el acceso a los datos y otras informaciones necesarias para realizar este estudio. Agradecen también la colaboración del Departamento de Gobierno y Empresas de la Universidad de los Lagos y del Departamento de Ingeniería Industrial y de Sistemas de la Pontificia Universidad Católica de Chile. También se agradece a los revisores anónimos y editores por sus sugerencias y comentarios que ayudaron a mejorar la presentación del artículo. Esta investigación contó con financiamiento del Programa de Mejoramiento de la Calidad y la Equidad de la Educación Superior (MECESUP ULA 2002).

REFERENCIAS

- Picard P. Auditing claims in the insurance market with fraud: the credibility issue. *J Public Econ*. 1996;63(1):27-56.
- Derrig R. Insurance fraud. *J Risk Insur*. 2002; 69(3):271-87.
- Major J, Riedinger D. EFD: a hybrid knowledge/statistical-based system for the detection of fraud. *J Risk Insur*. 2002;69(3):309-24.
- Brockett P, Derrig R, Golden L, Levine A, Alpert M. Fraud classification using principal component analysis of RIDITSs. *J Risk Insur*. 2002;69(3):341-71.
- He HX, Wang JC, Graco W, Hawkins S. Application of neural networks to detection of medical fraud. *Expert Syst Appl*. 1997;13(4): 329-36.
- Fraude a la seguridad social: cae una red empresarial que vendía contratos falsos a cambio de subsidios en Jaén. *El Mundo.es* (España) [publicación electrónica en Internet]. 2007 marzo 1. Hallado en <http://www.elmundo.es/elmundo/2007/03/01/espana/1172739110.html>. Acceso el 15 de septiembre de 2008.
- Santín D, Romay López R. El control del gasto público por incapacidad temporal mediante redes neuronales. *Hacienda Pública* (España). 2003;165:53-78. Hallado en http://www.ief.es/Publicaciones/Revistas/Hacienda%20Publica/165_ControlGasto.pdf. Acceso el 15 de septiembre de 2008.
- Baltussen B, Rhodes G, Narh-Bana SA, Agyepong I. Management of mutual health organizations in Ghana. *Trop Med Int Health*. 2006; 11(5):654-9.
- United Kingdom, National Audit Office. National insurance fund account 1998-1999. Report of the Comptroller and Auditor General. London: National Audit Office; 2000. Hallado en http://www.nao.org.uk/publications/nao_reports/9900146nif.pdf. Acceso el 15 de septiembre de 2008.
- United Kingdom, National Audit Office. Getting it right, putting it right. Improving decision making and appeals in the social security benefits. Report of the Comptroller and Auditor General. London: National Audit Office; 2000. Hallado en http://www.nao.org.uk/publications/nao_reports/02-03/02031142.pdf. Acceso el 15 de septiembre de 2008.
- Tokman M, Rodríguez J, Larraín F. Subsidios por incapacidad laboral 1991-2002: incentivos institucionales, crecimiento del gasto y una propuesta de racionalización. *Estudios Públicos*. 2004;93:219-47. Hallado en http://www.cepchile.cl/dms/lang_1/doc_3248.html. Acceso el 15 de septiembre de 2008.
- Bolton RJ, Hand DJ. Statistical fraud detection: a review. *Stat Sci*. 2002;17(3):235-49.
- Muñoz V, Pulido B. Cerca del 27% del gasto anual por licencias corresponde a fraudes. *Diario El Mercurio* (Chile). 2007 junio 4; sección Economía y Negocios. Hallado en <http://www.economiaynegocios.cl/noticias/noticias.asp?id=27361#>. Acceso el 15 de septiembre de 2008.
- Levy JP, Varela J. *Análisis multivariable para las ciencias sociales*. Madrid: Pearson Educación; 2003.
- United States of America, Congressional Budget Office. Administrative costs of private accounts social security. March, 2004. Washington, D.C.: Congress of the United States of America; 2004. Hallado en <http://www.cbo.gov/ftpdocs/52xx/doc5277/Report.pdf>. Acceso el 15 de septiembre de 2008.
- Baker D, Kar B. Defined contributions from workers, guaranteed benefits for bankers: the World Bank's approach to social security reform. Washington, D.C.: Center for Eco-

- conomic and Policy Research; 2002. Hallado en URL: <http://www.cepr.net/documents/publications/World-Bank-Social-Security.pdf>. Acceso el 15 de septiembre de 2008.
17. Departamento de Estadísticas e Información en Salud. Indicadores básicos de salud 2005. Santiago: Ministerio de Salud; 2006. Hallado en: <http://deis.minsal.cl/Base051.pdf>. Acceso el 15 de septiembre de 2008.
 18. ISAPRES. ISAPRES de Chile. Historia del ISAPRES [sitio en Internet]. Santiago: ISAPRES; sin año. Hallado en <http://www.isapre.cl/modulos.php?mod=phtml&fn=f4d3fce61827bf80098b989fd9c1c5b8>. Acceso el 15 de septiembre de 2008.
 19. Gobierno de Chile, Superintendencia de Salud. Boletín Estadístico 2006. Santiago: Superintendencia de Salud; 2007. Hallado en <http://www.supersalud.cl/documentacion/569/article-3584.html>. Acceso el 15 de septiembre de 2008.
 20. Gobierno de Chile, Superintendencia de Seguridad Social. Licencia médica [sitio en Internet]. Santiago: Superintendencia de Seguridad Social; sin fecha. Hallado en [http://www.suseso.cl/OpenDocs/asp/pagDefault.asp?boton=Doc203&argInstanciaId=203&argCarpetaId=306&argTreeNodosAbiertos=&argTreeNodoActual=306&argTreeNodoSel=306&argTreeNodosAbiertos=\(306\)](http://www.suseso.cl/OpenDocs/asp/pagDefault.asp?boton=Doc203&argInstanciaId=203&argCarpetaId=306&argTreeNodosAbiertos=&argTreeNodoActual=306&argTreeNodoSel=306&argTreeNodosAbiertos=(306)). Acceso el 15 de septiembre de 2008.
 21. Mesa F, Kaempffer AM. 30 años de estudio sobre ausentismo laboral en Chile: una perspectiva por tipos de empresas. *Rev Med Chile*. 2004;132(9):1100-8. Hallado en http://www.scielo.cl/scielo.php?pid=S0034-98872004000900012&script=sci_arttext. Acceso el 15 de septiembre de 2008.
 22. Ortega P, Figueroa C, Ruz G. A medical claim fraud/abuse detection system based on data mining: a case study in Chile. *International Conference on Data Mining*; 2006 June 26-29; Las Vegas, Nevada, USA. Hallado en <http://www.mec.cf.ac.uk/~scegr2/pub/DMI5560.pdf>. Acceso el 15 de septiembre de 2008.
 23. Hosmer DW, Lemeshow S. *Applied logistic regression*. 2nd ed. New York: John Wiley & Sons; 2000.
 24. Raftery AE. Bayesian model selection in social research (with discussion). In: Marsden PV, ed. *Sociological methodology*. Cambridge, Massachusetts: Blackwell; 1995. Pp. 111-96.

Manuscrito recibido el 26 de octubre de 2007. Aceptado para publicación, tras revisión, el 17 de junio de 2008.

ABSTRACT

Fraud in the health systems of Chile: a detection model

Objectives. To develop a model for detecting cases of organized fraud in Chile based on data from the legal forms for medically authorized leave (*formulario legal de licencia médica curativa*-MAL) and to establish the relevance of this data to fraud detection.

Methods. A binomial logistic regression model was employed using four variables from the MAL form, a national requirement for illness-related work absences: the number of legal absences taken by a single person, the number of days authorized by the prescribing doctor, the total cost per illness, and a dichotic variable reflecting whether or not the diagnosis is one that can be proven. The analysis involved 4 079 MAL forms that had been submitted in 2003 to a private health provider and of which 356 were already identified as fraudulent by a panel of medical fraud experts.

Results. The model successfully identified 99.71% of the fraudulent medical authorizations and 99.86% of the non-fraudulent, according to the criteria of the panel of fraud experts. Three of the variables employed had statistically-significant independent predictive power. The positive predictive value of the proposed model was 98.59%, while its negative predictive value was 99.97%.

Conclusions. The binomial logistic model that was developed uses four variables that are common to all MAL forms in use by Chile's public as well as private insurers, permitting separation of fraudulent from non-fraudulent requests to be more accurate, more timely, and at a cost lower than that of an expert panel.

Key words

Fraud; insurance, health; decision support techniques, Chile.