

# Respuesta inmediata ante grandes emergencias: modelo matemático basado en redes neuronales y árboles de decisión

JUAN ANTONIO SÁNCHEZ LOSADA<sup>1</sup>, SARA ROMERO SÁNCHEZ<sup>1</sup>, MERCEDES CAAMIÑA GARCÍA<sup>1</sup>, NELLY HABED LOBOS<sup>1</sup>, JUAN FEDERICO JIMÉNEZ CARRASCOSA<sup>1</sup>, BERNARDO TOUZA GARMA<sup>1</sup>, ÁNGEL MIGUEL GIL GONZÁLEZ<sup>2</sup>, PEDRO SANZ MATA<sup>2</sup>, JONATAN RAMIRO MOYANO<sup>2</sup>

<sup>1</sup>Médico de Emergencias, Servicio de Salud de Castilla-La Mancha (SESCAM), Toledo, España. <sup>2</sup>Técnico de Sistemas y Comunicaciones SESCAM, Toledo, España.

## CORRESPONDENCIA:

Juan Antonio Sánchez Losada  
Apartado de Correos 486  
45080 Toledo  
E-mail: urgencias@gmail.com

## FECHA DE RECEPCIÓN:

24-6-2011

## FECHA DE ACEPTACIÓN:

20-7-2011

## CONFLICTO DE INTERESES:

No existe conflicto de intereses.

## AGRADECIMIENTOS:

Este proyecto ha sido becado por el Área de Investigación de la Fundación Sociosanitaria de Castilla-La Mancha.

**Objetivo:** En la gestión de recursos ante incidentes con múltiples víctimas (IMV), el factor incertidumbre en un entorno de caos inicial se convierte en el peor enemigo del decisor. El objetivo fue diseñar un modelo matemático predictivo en este entorno de decisiones para mejorar la fase inicial de la gestión de recursos ante una gran emergencia.

**Método:** Se partió de una base de datos de incidentes reales con múltiples víctimas en Castilla-La Mancha en los últimos cinco años, sobre la que se trabajó con 10 variables categorizadas en función de su peso en la gravedad de la emergencia. Se diseñó una red neuronal, que aprende sobre los casos reales, y por otro lado se generó un árbol de decisión con el fin de buscar la mejor respuesta entre ambos modelos. Se ha encontrado una importante limitación metodológica, ya que prácticamente todos los IMV analizados son accidentes de tráfico.

**Resultados:** El modelo basado en árboles de decisión da más información y mayor variabilidad que la red neuronal e identifica 6 grupos homogéneos en función del "número de afectados iniciales", "tipología del incidente" y "entorno" ( $p < 0,05$ ).

**Conclusiones:** Es posible crear un modelo matemático predictivo con las variables consideradas que permite una mejor gestión de recursos ante un IMV, pero se necesita un mayor número de casos reales pasados y con tipología más diversa, para ser aplicadas a casos reales futuros con la metodología propuesta. [Emergencias 2012;24:175-180]

**Palabras clave:** Servicios de emergencias sanitarias. *Triage*. Redes neuronales artificiales. Catástrofes. Transporte sanitario.

## Introducción

Con carácter general, por lo inesperado de los accidentes con múltiples víctimas, grandes emergencias y catástrofes, el tratamiento de lesionados se lleva a cabo en condiciones iniciales de escasez de profesionales y equipamientos, especialmente en entornos de gran dispersión geográfica. Por ello es esencial disponer de herramientas que mejoren el uso óptimo de los recursos disponibles<sup>1</sup>.

Los modelos matemáticos aplicados a la gestión de las catástrofes no sólo proporcionan métodos para ganar mayor comprensión de los procesos inherentes a las catástrofes, sino que también dan instrumentos con los cuales controlar algunos aspectos de estos procesos. Estos modelos tratan de

evitar que estas situaciones, por inesperadas, sean imprevistas. Destacan las aportaciones de De Bøer quien concluye, tras analizar más de 400 grandes emergencias, que existe relación matemáticamente objetivable entre la gravedad del incidente y los porcentajes de lesionados muy graves, graves y leves que previsiblemente se van a producir<sup>2</sup>.

La gravedad de un accidente múltiple, gran emergencia o catástrofe se establece fundamentalmente a través del índice global de gravedad clínica de los afectados. Los grupos de víctimas fundamentales a considerar son los "muy graves", T1, y los "graves", T2 que precisan soporte vital in situ, transporte y hospitalización pues cualquier retraso en su atención con los recursos adecuados supondría un incremento en el número de falleci-

dos, de forma que la gravedad global del incidente viene dada por el cociente:  $T1 + T2/T3^3$ .

El presente trabajo trata de aproximarse a una herramienta que ayude en el proceso de toma de decisiones cuando se produce la gran emergencia, desde el convencimiento de que una adecuada reorganización y unas decisiones iniciales consecuentes proporcionan a los heridos una mayor oportunidad de supervivencia, y de que uno de los elementos necesarios en la toma de decisiones ante múltiples víctimas es una estimación general de la gravedad esperable, así como los requerimientos previsibles de traslado y hospitalización derivados de dicha estimación. De la famosa curva de Trunkkey (distribución trimodal de trauma-muertes relacionada con el tiempo)<sup>4</sup>, se desprende que el segundo y tercer picos de las muertes por traumatismos –alrededor del 50% de las muertes– podrían ser prevenidos mediante una adecuada respuesta prehospitalaria e intrahospitalaria, precoz y adaptada a las necesidades. Cualquier acción dirigida a mejorar la atención precoz de estos grupos de víctimas con los recursos adecuados redundaría en un aumento de la supervivencia.

El objetivo general del presente trabajo es diseñar dos modelos matemáticos de ayuda a la toma de decisiones ante grandes emergencias en cuanto a la gravedad global de las víctimas y las necesidades de recursos: uno basado en el aprendizaje de una red neuronal artificial (RNA) sobre grandes emergencias pasadas, otro basado en árboles de decisión, y comparar sus rendimientos en cuanto a predicción.

## Método

Para dar respuesta a las preguntas generadas por el objetivo del presente trabajo, se han desarrollado las siguientes fases descritas a continuación:

Fase 1. Análisis retrospectivo de incidentes con múltiples víctimas (IMV) en Castilla-La Mancha (España), una comunidad autónoma con alta dispersión de población (26 hab/km<sup>2</sup>). Los datos del estudio se tomaron de las bases de datos del centro de atención de emergencias 1-1-2 (176 casos y 2.755 lesionados), que han servido para entrenar la RNA y obtener el árbol de agregación. Los criterios de inclusión han sido: incidentes con 5 o más afectados y comprendidos entre los años 2006 y 2010. Los criterios de exclusión han sido: incidentes con múltiples afectados en los que no se haya precisado la intervención de recursos sanitarios o incidentes que finalmente hayan sido finalizados como “nulos” o “anulados”.

Se ha encontrado una importante limitación metodológica, ya que los IMV obtenidos en la base de

datos analizada son casi todos accidentes de tráfico.

La variable principal del proyecto fue extraer un patrón del comportamiento inicial de una gran emergencia a través de un modelo matemático apoyado en una RNA y de un árbol de decisión<sup>5</sup>.

Para ello se considerarán 10 variables como insumos que van a influir en la gravedad en función del valor asignado a cada una de las variables, a las que se asigna una gravedad parcial (Tablas 1 a 3). La suma de ellas ofrece la gravedad global del incidente: tipología, número inicial de afectados, horario en que se produce, fallecimientos confirmados iniciales, gran alarma generada, meteorología adversa, tipo de entorno o escenario, área del incidente, duración prevista y número de focos. La variable dependiente considerada es la suma de graves + muy graves ( $T_2 + T_1$ ), que es el numerador de la ecuación de gravedad ( $T_1 + T_2/T_3$ ).

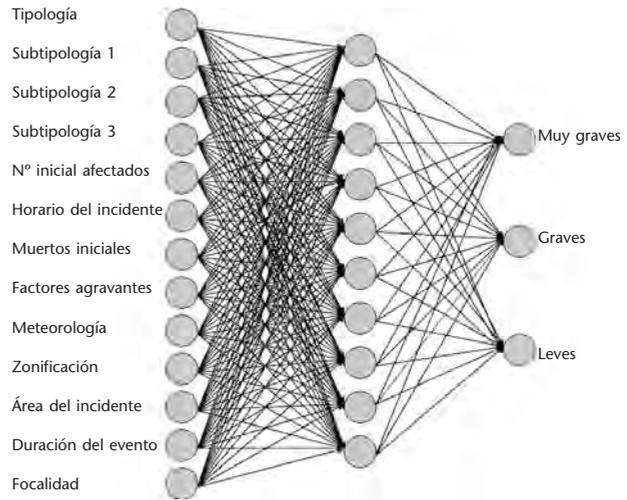
Fase 2. Las RNA son conceptos matemáticos, inspirados en las redes neuronales biológicas, que pueden interactuar con el mundo real. Estas redes “aprenden” y de hecho, antes de utilizarlas con garantías, hay que entrenarlas y, tras un determinado número de ensayos, la red es capaz de discernir, basándose en patrones obtenidos a través de series de casos suficientes de bases de datos históricas. Cada neurona está conectada con otras neuronas mediante enlaces de comunicación, cada uno de los cuales tiene asociado un peso. En los pesos se encuentra el conocimiento que tiene la RNA acerca de un determinado problema<sup>6,7</sup>. El modelo de RNA empleado se basa en una arquitectura del tipo Perceptrón Multicapa entrenada mediante la regla de aprendizaje *backpropagation* (propagación del error hacia atrás). El Perceptrón Multicapa está formado por una capa de entrada, una capa de salida y una o más capas ocultas o intermedias; la información se transmite desde la capa de entrada hasta la capa de salida y cada neurona está conectada con todas las neuronas de la siguiente capa. El número de neuronas de la capa de entrada está determinado por el número de variables predictoras. Cada neurona de entrada tiene como misión recibir y transmitir a la siguiente capa el valor de una de estas variables. Finalmente se consideraron 13 neuronas, ya que la variable tipología podía tener diferentes valores para un mismo incidente (por ejemplo: un atentado terrorista con bombas puede provocar un accidente aéreo), por lo que se ofreció a la red una tipología principal y 3 subtipos relevantes en la gravedad del incidente. No existe una receta que indique el número óptimo de neuronas en la capa oculta y se debe usar el mínimo número de neuronas ocultas con las cuales la red rinda de forma adecuada,

**Tabla 1.** Tipología del Incidente (incluye subtipos)

Variable 1		Gravedad
Accidente de tráfico	Meteorológicos	0
Polígono industrial	Viviendas unifamiliares	
Alta velocidad: autovías, AP	Fuga gas no controlada	1
Atrapados	Accidente tren de pasajeros	
Autobuses implicados	Accidente biológico	
Inundaciones	Accidente en eventos	
Grandes retenciones	Atentado terrorista	
Incendio o explosión	Avalanchas, alud	
Camiones involucrados	Colisión meteorito	
Lluvias intensas, nevadas	Derrumbes, explosiones	
Mercancías peligrosas	Incendios	
Oficinas, bancos, edificios públicos	Intoxicación masiva	
Polideportivo, estadio	Epidemia	
Puente autopista	Plaza de toros	2
Teatro, cine, discoteca	Químicos. Nube tóxica	
Camping	Refinería	
Centro comercial, hotel	Avión de pasajeros	
Difícil acceso	Accidente industrial	
Ed. viviendas, colegio, hospital	Accidente nuclear	
Residencia de ancianos	Disturbios	
Lluvias torrenciales, ciclón	Erupción volcánica	
Grandes nevadas, huracán	Huracán	
Riesgo de bombas, armas	Naufragio barco pasajeros	
Túnel, metro	Terremoto	

Ed.: edificio.

y esta red funcionó de forma óptima con 10 neuronas ocultas (Figura 1). Y por último, el número de neuronas de la capa de salida está determinado por el número de categorías o clases que tiene el problema. En el presente trabajo, las neuronas de salida darán como resultado los porcentajes de lesionados en función de su previsible gravedad y necesidades de recursos (T1, T2 y T3). Existen numerosos sistemas de clasificación de víctimas en función del número de categorías consideradas atendiendo a la gravedad y las posibilidades de supervivencia, y en este estudio se ha considerado inicialmente el tripolar (T1, T2, T3) para obtener la máxima información de la base de datos de lesiones reales a través del código CIE9mc (clasificación internacional de enfermedades, 9ª revisión modificación clínica), aunque posteriormente los muy graves (T1) y los graves (T2) se han agrupado para la comparación entre modelos, por varios motivos: la sencillez del método bipolar, la rapidez que imprime en la toma de decisiones, la información incompleta inicial ante un incidente de estas características y el objetivo fundamental de



**Figura 1.** Red neuronal artificial utilizada en el grupo de estudio.

esta clasificación, dirigido hacia las decisiones sobre recursos iniciales, más que sobre los propios pacientes. Así pues, los T1 y T2 requerirán medidas de soporte vital básico y/o avanzado inmediato e ingreso urgente en el hospital, mientras que los T3 (leves) pueden ser derivados a centros de atención primaria con o sin posibilidad de diagnóstico radiológico según el tipo de lesiones<sup>8</sup>.

Fase 3. El segundo modelo matemático planteado trata de obtener la relación entre variables a través de la segmentación progresiva de la muestra para obtener un árbol de agregación. Esta técnica nos permite conocer cuáles son las variables Xi que mejor explican o predicen Y. En las técnicas estadísticas clásicas esta relación se expresa mediante una ecuación, mientras que aquí se expresa mediante un árbol de agregación. Para el análisis de datos se ha utilizado el paquete estadístico SPSS 17.0. y se ha realizado además un árbol de agregación y un corte transversal del mismo, que nos ha dado 6 agrupaciones de variables con un comportamiento significativamente homogéneo; para ello se ha utilizado la herramienta SPAD.N, 2005. ("système portable pour l'analyse des données". Ver. 5.6) DECISIA, Francia. Una vez obtenido el árbol, en cada nodo interior se selecciona la "mejor variable" en cuanto que

**Tabla 2.** Variables 2 a 5 (número inicial de afectados, horario del incidente, fallecimientos iniciales comunicados, gran alarma en la población)

Variable 2: Número inicial de afectados	Variable 3: Horario del incidente	Variable 4: fallecimientos iniciales	Variable 5: Gran alarma en la población	Gravedad
< 10	Nocturno en accidentes y diurno (9:01-17:00)	No	No objetivada	0
11-50	Nocturno en viviendas (20:01-7:00)	Sí	Sí (> 40 llamadas de diferentes alertantes)	1
> 100	Hora punta-tarde (17:01-20:00), operaciones salida y hora punta-mañana (7:01-9:00)	Inicial >10	> 100 llamadas	2

**Tabla 3.** Variables 6 a 10 (meteorología adversa, entorno, área del incidente, focos, duración previsible del incidente)

Variable 6: meteorología adversa	Variable 7: entorno	Variable 8: área donde aparezcan víctimas	Variable 9: focos	Variable 10: duración incidente	Gravedad
No consta	Urbano/rural próximo (hospital en la misma localidad/o a menos de 30 minutos)	< 1 km	1 foco	< 1 hora	0
Lluvias torrenciales, nevadas	Rural alejado (hospital a más de 30 minutos)	1-10 km	2 focos	1-24 horas	1
Grandes nevadas, huracán, ciclón	Desfavorable (extrema pobreza, situación de guerra, accesos cortados, escasez de medios, huracán o nube tóxica o terremoto en zona urbana que afecte a los propios recursos de emergencia)	> 10 km	> 2 focos	> 24 horas	2

proporciona más información sobre Y. Cada nodo del árbol es un segmento de la muestra considerada. Se considera en el nodo inicial o raíz el conjunto de todos los individuos. Se examinan todas las variables consideradas y se generan nodos descendientes del inicial con los segmentos de individuos definidos por la subdivisión seleccionada. Se detiene el procedimiento cuando todos los segmentos se han declarado terminales.

## Resultados

Se observó que la mayoría de los casos de la muestra, que es la población total de incidentes de estas características en el periodo considerado, son accidentes de tráfico con similar gravedad global; y con la predicción de la RNA se obtiene una estimación media a la salida de 53,31% de casos leves y 46,69% de casos graves y muy graves. En cambio los valores medios reales observados presentan una media de 74,1% y 25,9% respectivamente.

El árbol de agregación obtenido (Figura 2) elige las variables por orden de relevancia en la predicción de la variable dependiente considerada. Con una  $p < 0,05$ , produce subdivisiones sucesivas de la muestra, que presentan diferencias de medias en el porcentaje de T1 + T2. La primera variable elegida por el árbol fue el número de afectados, y a mayor número, el porcentaje de T1 + T2 (graves + muy graves) disminuye. Esta variable segmenta la muestra en dos grupos: el primer segmento corresponde a dos categorías de afectados: "número de afectados menor de 10", con 90 incidencias, donde el porcentaje de graves-muy graves es 35,0%; y el segundo segmento, con 86 incidencias, corresponde a las categorías de "más de 10 afectados", y el porcentaje de gravedad es un 15,5%. En un segundo nivel se subdividen las 90 incidencias del nodo 1 según los valores de SUBTIPO1 en primer lugar, y destaca el ENTORNO a continuación, con un pronóstico del 46,7% de T1 + T2 en un 16,5% de los incidentes en entorno rural. De esta forma el árbol

determina en total 6 tipos, clases, segmentos de incidencias (nodos 2, 5, 6, 7, 9 y 10). Los valores medios de T1 + T2 van desde el 15,5% al 46,7%.

## Discusión

En el presente estudio el modelo basado en redes neuronales se aleja de los valores reales en su predicción de T1 + T2 (real: 25,9%, predicción: 46,7%); mientras que el árbol de agregación no da una predicción única, sino que divide la muestra en 6 grupos en función de las diferentes variables, otorgando una mayor variabilidad de resultados.

La variable "número de afectados" es la primera seleccionada en el árbol de decisión con un comportamiento inverso en la relación a la variable dependiente, fácilmente explicable, ya que la mayor gravedad de lesiones se produce en la zona de impacto (choque inicial, explosión, etc.), y a medida que nos alejamos del foco, aunque tengamos un mayor número de víctimas, la gravedad es menor (choques en cadena, daños secundarios a distancia en explosiones, incendios, etc.). Las variables "entorno" y "duración del incidente" le siguen en importancia predictiva sobre el porcentaje final de graves y muy graves, de forma que a mayor complejidad de la atención (entorno rural vs urbano) y mayor duración de la misma (superior a una hora) aumenta el porcentaje de graves-muy graves y por tanto la gravedad del incidente en términos de gravedad de lesiones y necesidad de recursos avanzados: un 46,7% para incidentes de menos de 10 afectados en entorno rural (que aglutinan al 16,5% de los incidentes), frente a un 15,5% de graves-muy graves en incidentes de más de 10 afectados (que aglutinan al 48,9% de IMV analizados).

Las RNA han sido aplicadas de forma satisfactoria en la predicción de diversos problemas en diferentes áreas de conocimiento –biología, medicina, economía, ingeniería, psicología, etc.– y han obtenido excelentes resultados respecto a los modelos derivados de la estadística clásica. La virtud de las RNA reside en su capacidad para aprender

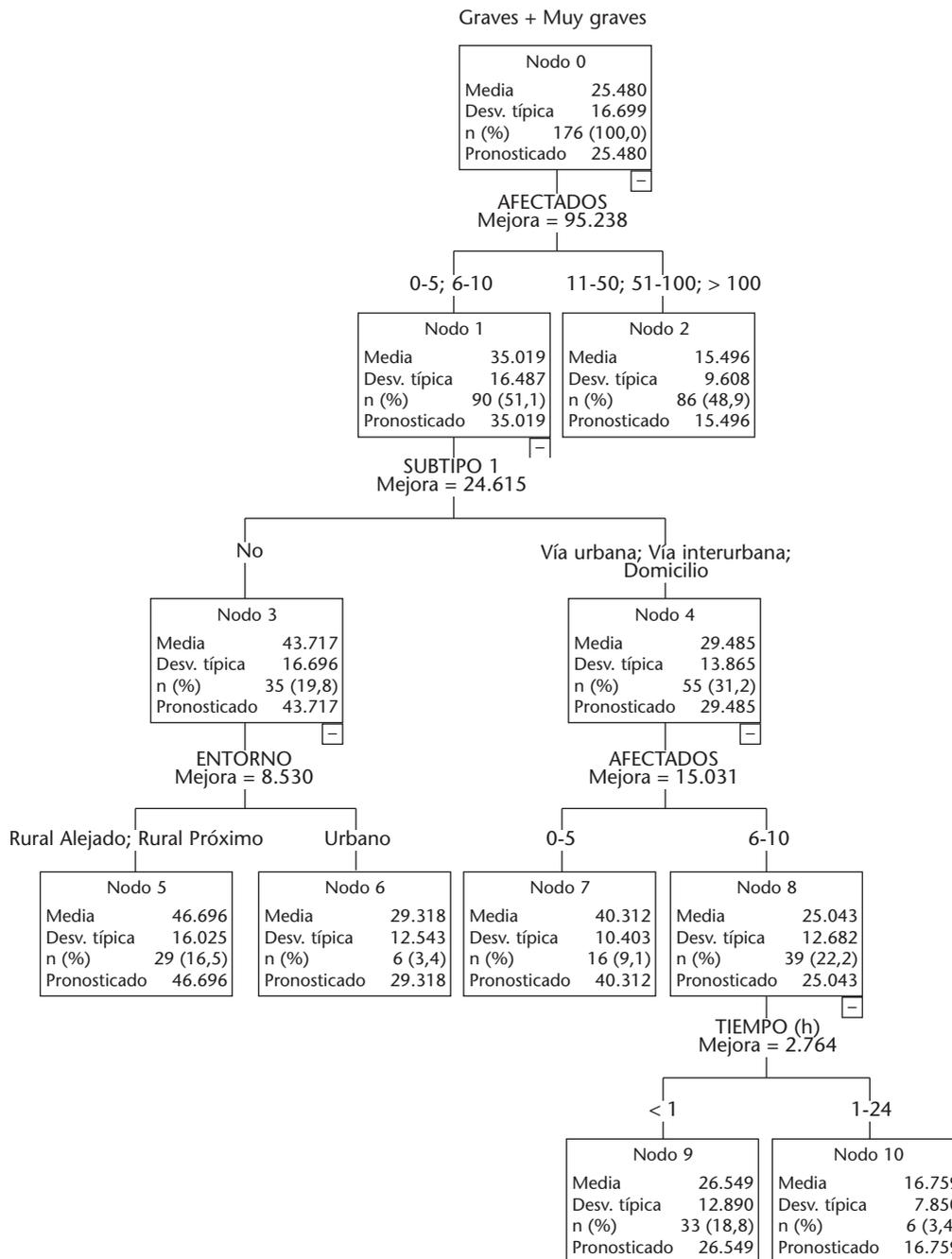


Figura 2. Árbol de agregación del grupo de estudio.

funciones complejas o no lineales entre variables sin necesidad de imponer presupuestos o restricciones de partida en los datos. No obstante, precisan una información variada para dar respuestas variadas. En el presente trabajo, la escasa variedad de tipología de los IMV analizados (casi todos accidentes de tráfico) resta eficacia a la red obtenida. El árbol de agregación, como segunda técnica empleada sí que aporta resultados equiparables a los reales en función de la importancia de cada una de las variables consideradas, aunque no es

una herramienta tan flexible como la red neuronal que permita su aplicación práctica ante IMV futuros por sí sola.

Con esta información se extraen las siguientes conclusiones:

Primera. Es posible realizar un modelo predictivo de la gravedad de los IMV, a través de un análisis de incidentes reales con las variables consideradas en este estudio.

Segunda. La muestra considerada de 176 casos en Castilla-La Mancha es suficiente para obtener

información del peso de las variables consideradas, pero insuficiente para extraer un modelo predictivo aplicado basado en redes neuronales por lo homogéneo de la tipología (prácticamente la totalidad son accidentes de tráfico). El modelo de red neuronal diseñado debe ser entrenado con incidentes de mayor complejidad y más variada tipología que los aquí empleados.

Tercera. El modelo basado en árboles de agregación es el que más información aporta sobre el papel de cada variable considerada en la gravedad total del incidente, e identifica 6 grupos homogéneos con porcentajes de heridos graves y muy graves que oscilan entre el 15,5 y el 46,7%, en función del número de afectados iniciales, tipología del incidente, el entorno rural o urbano y la duración del incidente menor o mayor de una hora. El árbol de agregación tendría teóricamente posibilidades predictivas reales en las condiciones consideradas, ya que permite evidenciar las relaciones existentes entre variables y la importancia relativa de cada una de ellas. El valor real está comprendido en la horquilla predictiva del árbol, pero no es una herramienta lo suficientemente flexible para ser aplicada en la práctica.

Cuarta. Si bien el modelo de árbol de decisión define mejor el comportamiento de los índices de gravedad según las variables consideradas sobre la muestra de este trabajo, seguimos defendiendo la flexibilidad que ofrece la red neuronal al aprender de incidentes futuros que implementados en la red son capaces de alterar los pesos de las variables y mejorar la predicción. Se necesitan por tan-

to más casos reales y más variados. Y el árbol de agregación pasaría a ser una herramienta complementaria a la red, ya que nos permite rediseñar y mejorar la red a través del análisis profundo que hace de las variables que son relevantes y las que apenas influyen en los resultados finales.

Quinta. Debería ser considerada la recogida homogénea de información tras una gran emergencia en registros específicos, y la opción de crear un registro común compartido por los diferentes servicios de emergencias, que permitiera compartir experiencias y lecciones aprendidas, así como alimentar estas herramientas predictivas para mejorar nuestra atención a estos incidentes.

## Bibliografía

- 1 Sánchez Losada JA, Caamiña García M, Jiménez Carrascosa JF, Touza Garma B, Romero Sánchez S, Gregorio Sanz M, et al. Modelo de respuesta inmediata ante grandes emergencias para la atención sanitaria en entornos de dispersión geográfica. Puesta al día en urgencias, emergencias y catástrofes. 2009;9:78-88.
- 2 De Boer J. Order in Chaos: modellin, medical management in disaster. Eur J Emerg Med. 1999;6:141-8.
- 3 Álvarez Leiva C, Macías Seda J. Triage: generalidades. Puesta al día en Urgencias, Emergencias y Catástrofes. 2001;2:125-33.
- 4 Trunkey DD. Trauma Sci Am. 1983;249:20-7.
- 5 Baker SP, O'Neill B, Haddon W, Long WB. The injury severity score: a method for describing patients with multiple injuries and evaluating emergency care. J Trauma. 1974;14:187-96.
- 6 Hilerá JR, Martínez VJ. Redes Neuronales Artificiales: fundamentos, modelos y aplicaciones. Madrid: RA-MA Editorial; 1995.
- 7 Mussone L, Ferrari A, Oneta M. An analysis of urban collisions using an artificial intelligence model. Accident Analysis and Prevention. 1999;31:705-18.
- 8 Meislin H, Criss EA, Judkins D, Berger R, Conroy C, Parks B, et al. Fatal trauma: the modal distribution of time to death is a function of patient demographics and regional resources. J Trauma. 1997;43:433-40.

## Rapid response to large-scale emergencies: a neural network model and a decision-tree algorithm

Sánchez Losada JA, Romero Sánchez S, Caamiña García M, Habed Lobos N, Jiménez Carrascosa JF, Touza Garma B, Gil González AM, Sanz Mata P, Ramiro Moyano J

**Objective:** The greatest challenge to decision-making during the management of emergencies with multiple victims is uncertainty in an initially chaotic environment. Our objective was to develop a predictive model to improve response and early resource management in the early-phase environment of chaotic uncertainty during large-scale emergencies.

**Methods:** A database of information on real incidents with multiple victims in Castile-La Mancha, Spain, in the last 5 years was used to study the weight of 10 categorical variables and their effect on the seriousness of the emergencies. A neural network was designed to learn about these real cases, and a decision tree was generated, to study which of the 2 approaches gave the best results. An important design limitation was that nearly all the incidents analyzed involved traffic accidents.

**Results:** The model based on decision-tree analysis gave more information and greater variability. It proved superior to the neural network, identifying 6 homogeneous groups according to the following factors: number of initial victims, type of incident, and environment ( $P < .05$ ).

**Conclusions:** A predictive model can be based on the considered variables in the interest of improving resource management during a large-scale emergency. However, development based on a larger number of real incidents of different types would be needed before such a model could be applied during real future incidents. [Emergencias 2012;24:175-180]

**Key words:** Emergency health services. Triage. Neural networks (computers). Catastrophes. Ambulance services.