

Modelo basado en lógica borrosa para estratificar el riesgo de la cirugía cardíaca

Fuzzy Logic-Based Model to Stratify Cardiac Surgery Risk

RAÚL A. BORRACCI^{1,2,3,4}, EDUARDO B. ARRIBALZAGA

RESUMEN

Introducción: La práctica clínica se desarrolla en un contexto de información incierta, en el que el conocimiento "experto" ha demostrado que es muy eficiente para la toma de decisiones.

Objetivo: Desarrollar y validar un modelo basado en lógica difusa o borrosa para predecir el riesgo de mortalidad en cirugía cardíaca.

Material y métodos: Se incorporaron prospectivamente 450 pacientes sometidos a cirugía cardíaca y se cotejó la predicción de riesgo de mortalidad en base a cinco puntajes: 1) la opinión de un "experto", 2) el resultado de un sistema basado en lógica difusa según el conocimiento experto, 3) Parsonnet, 4) Ontario y 5) EuroSCORE. El modelo de lógica difusa se desarrolló en las siguientes etapas: selección por un experto de las variables predictivas de mortalidad, confección de tablas de influencia entre variables, construcción de un mapa cognitivo borroso (MCB) e implementación en una red neuronal artificial, determinación por el experto del puntaje de riesgo por paciente, cálculo del riesgo del conjunto de prueba según los predictores borrosos, determinación del riesgo del conjunto de validación usando el MCB ya calibrado y comparación de los resultados con los otros modelos según concordancia y precisión con curvas ROC.

Resultados: El modelo calibrado se usó para predecir los resultados del conjunto de validación (360 pacientes), a quienes se les determinó el puntaje del MCB y los riesgos pronosticados por Parsonnet, Ontario y EuroSCORE. Las áreas ROC demostraron que el MCB tuvo por lo menos el mismo desempeño para predecir mortalidad (ROC = 0,793 vs. 0,775, 0,767, 0,741 y 0,701 para el EuroSCORE, "experto", Ontario y Parsonnet, respectivamente).

Conclusiones: Se propone un sistema capaz de aprovechar el conocimiento experto mediante el uso de lógica difusa y la implementación de un sistema experto para predecir la mortalidad en cirugía cardíaca. El modelo no solo imitó los resultados obtenidos por el "experto", sino que también tuvo el mismo desempeño que otros puntajes.

Palabras clave: Procedimientos quirúrgicos cardíacos - Pronóstico - Medición de riesgo - Lógica difusa

ABSTRACT

Background: Medical practice is usually performed in a context of uncertainty, where expert knowledge has shown to be efficient in the decision-making process.

Objective: The aim of this study was to develop and validate a fuzzy logic-based model to predict cardiac surgery mortality risk.

Methods: Four hundred and fifty patients undergoing cardiac surgery were prospectively included in the study and mortality risk was predicted based on five scores: 1) "clinical expert" opinion, 2) fuzzy logic-based system according to expert knowledge, 3) Parsonnet, 4) Ontario and 5) EuroSCORE. The fuzzy logic model was developed in the following stages: expert selection of different mortality predictive variables, tables of influence among variables, construction of a fuzzy cognitive map (FCM) and its implementation in an artificial neuronal network, expert-determined patient risk score, test set risk calculation based on fuzzy predictors, validation set risk using calibrated FCM, and comparison with the other scores according to the level of agreement and precision with ROC curves.

Results: The calibrated model was used to predict the outcome of the validation set (360 patients), based on the FCM score and risk predicted by Parsonnet, Ontario and EuroSCORE. The ROC areas showed that FCM had at least the same performance as other scores to predict mortality (ROC=0.793 vs. 0.775, 0.767, 0.741 and 0.701 for EuroSCORE, "expert", Ontario and Parsonnet, respectively).

Conclusions: A fuzzy logic-based system employing expert knowledge and the implementation of an expert system is postulated to predict cardiac surgery mortality risk. The model not only mimicked the outcomes obtained by the "expert", but had the same performance as others risk scores.

Key words: Cardiac Surgical Procedures - Prognosis - Risk Assessment - Fuzzy Logic

REV ARGENT CARDIOL 2015;83:305-313. <http://dx.doi.org/10.7775/rae.v83.i4.6730>

Recibido: 04/06/2015 - Aceptado: 01/07/2015

Dirección para correspondencia: Raúl A. Borracci - La Pampa 3030, 1.° B - (1428) Buenos Aires, Argentina - e-mail: raborracci@gmail.com

Hospital de Clínicas, Facultad de Medicina, Universidad de Buenos Aires, Buenos Aires, Argentina

Bioestadística, Facultad de Ciencias Biomédicas, Universidad Austral, Pilar, Provincia de Buenos Aires, Argentina

⁴ Miembro Titular de la Sociedad Argentina de Cardiología

INTRODUCCIÓN

Los puntajes o scores de riesgo han sido ampliamente desarrollados en el ámbito de la cirugía cardíaca y continúan aún en crecimiento. (1-6) La metodología habitual para la confección de estos modelos es el tratamiento de datos masivos por medio de algún tipo de análisis multivariante o análisis bayesiano para determinar la ocurrencia del evento "muerte posoperatoria" asociada con un conjunto de variables predictoras independientes. (7-14) El proceso de toma de decisiones en la práctica médica se basa en gran medida en la interpretación de un conjunto de datos incompletos, cuya recolección implica también un margen de error. La acumulación de información clínica se obtiene directamente a partir de las percepciones del médico-experto, o a través de percepciones previas de situaciones similares, relacionadas a su vez con lo que se sabe de la enfermedad. Pero la falta de completitud y la inexactitud del conocimiento sobre esta situación clínica es lo que produce la incertidumbre bajo la cual se deberán tomar las decisiones. Finalmente, la incapacidad para listar el conjunto completo de antecedentes o consecuencias de una patología determinada o de cierta intervención sobre ella, así como la imposibilidad de evaluar todas las variables en juego que rodean muchas situaciones clínicas hacen que el juicio de un experto continúe teniendo una validez relevante al tomar decisiones en medicina.

Todos los puntajes de riesgo desarrollados hasta la actualidad se basan en la lógica *booleana* (verdadero / falso) que solo permite elegir entre dos o más posibilidades mutuamente excluyentes: un paciente puede ser o no diabético, puede tener o no una lesión del tronco de la coronaria izquierda, su función ventricular puede ser mejor o peor, condiciones que generarán un peso mayor o menor en la puntuación final del puntaje. Una vez elegido el valor de verdad de cada variable, el conjunto de datos se trata con rigurosas técnicas de análisis multivariante a fin de obtener una función capaz de predecir con cierto margen de probabilidad el resultado quirúrgico. Una hipótesis inicial podría suponer que, en parte, la deficiencia o dificultad para realizar una predicción correcta con estos puntajes se debería a una adjudicación incorrecta del valor de verdad o no de cada variable. Y aunque se pudieran definir todas las variables en función de su valor de verdad o no, quedarían aún otras situaciones que no pueden evaluarse con la lógica *booleana* tradicional: la edad de un paciente puede medirse en años, o en su condición de joven, viejo o geronte, pero ¿cómo podría medirse el aspecto de un paciente que parece más joven o viejo que su edad cronológica? El problema planteado hasta este punto señala las limitaciones de la lógica tradicional para resolver la situación de la predicción clínica. (15) A principios del siglo xx, J. Lukasiewicz desarrolló los principios de la lógica multivaluada, cuyos enunciados pueden tener valores de verdad comprendidos en una escala continua entre 0 (falso) y 1 (verdadero) de la lógica *booleana* clásica. Por ejemplo,

el enunciado "el vaso está lleno" en lógica tradicional tendría el valor de verdad 1 (verdadero) si el recipiente contiene tanto líquido como su capacidad máxima admite; por el contrario, si el vaso contiene el 90% de su capacidad total, el enunciado sería falso, con valor de verdad 0. La lógica multivaluada permite asignar diferentes grados de certeza; de esta forma, el valor de verdad sería 0,9 (casi verdadero). En 1965, Zadeh (16) introdujo el término lógica borrosa o difusa y desarrolló un álgebra completa para el tratamiento de estos conjuntos borrosos. La denominada lógica borrosa permite entonces tratar información imprecisa, como estatura media, temperatura baja o mucha fuerza, en términos de conjuntos difusos o imprecisos. Estas descripciones lingüísticas usadas tan frecuentemente en medicina clínica son especialmente aptas para su tratamiento con este método. (17-25) Por otra parte, los modelos basados en lógica difusa pueden implementarse en sistemas conocidos como mapas cognitivos borrosos. (26-28) Este mapa causal que representa el conocimiento del "experto" puede a su vez implementarse sobre una red neuronal artificial. (29) En resumen, la lógica borrosa es una clase de lógica multivaluada que permite derivar conclusiones a partir de información vaga, poco precisa o incluso ambigua. En muchos aspectos, la lógica borrosa imita a los procesos decisorios humanos, en los cuales se deben adoptar conductas o decisiones basadas en datos aproximados obtenidos a partir de sus sentidos.

En base al marco teórico expuesto hasta aquí, se definieron los siguientes objetivos del trabajo:

- Desarrollar y validar en forma prospectiva un modelo para la predicción del riesgo de mortalidad poscirugía cardíaca, basado en lógica borrosa.
- Implementar el modelo en un sistema artificial de redes neuronales en base a un mapa cognitivo borroso, capaz de funcionar como un sistema experto.
- Comparar los resultados de predicción obtenidos con este modelo contra diferentes sistemas de ajuste de riesgo validados internacionalmente.
- Analizar el grado de precisión de la predicción de cada modelo en base a los resultados de mortalidad observados.

MATERIAL Y MÉTODOS

Se incluyeron todos los pacientes adultos sometidos a cirugía cardíaca desde febrero de 2007 a marzo de 2008. Solo se excluyeron los pacientes a los que en el preoperatorio no se les pudo asignar alguno de los puntajes de riesgo elegidos o que no pudieron ser evaluados por el "experto" a fin de otorgarles un nivel de riesgo preoperatorio.

El diseño fue de tipo prospectivo, longitudinal y comparativo en el que se cotejaron la predicción de riesgo de mortalidad poscirugía cardíaca de distintos modelos de asignación de puntaje de riesgo. A cada uno de los pacientes quirúrgicos se le asignaron cinco niveles diferentes de riesgo de acuerdo con la opinión de un "experto" (R.A.B.), el resultado de un sistema automatizado desarrollado a partir del conocimiento experto y tres puntajes validados internacionalmente y de uso corriente al momento del estudio (Parsonnet, Ontario y

EuroSCORE). (7-9) En la Figura 1 se muestra el diagrama del diseño.

Secuencia de desarrollo del trabajo

Paso 1. Selección por un experto de las variables difusas predictivas de mortalidad, expresadas como variables nominales u ordinales, con sus correspondientes categorías (véase Tabla en Material suplementario en la web).

Paso 2. Confección de las tablas de influencia entre las variables y de estas con el punto final mortalidad. Para ello se usó un modelo estándar a partir del cual se aplicaron y ajustaron los "pesos" o "influencias" entre las variables, en una escala continua entre 0 (cero) y 1 (uno) (no afecta: 0,0; afecta ligeramente: 0,3; afecta moderadamente: 0,6 y afecta mucho: 1,0).

Paso 3. Construcción del mapa cognitivo borroso para predecir los resultados de la cirugía cardíaca en base a las variables borrosas preoperatorias. Este diagrama causal o mapa de conocimientos se construyó asignando a una serie de conexiones entre los nodos que mostraron la relación causal o la influencia entre las variables. Las conexiones entre los nodos llevan el valor asociado en el rango [0-1] de acuerdo con lo determinado en las tablas de influencia.

Paso 4. La implementación del mapa cognitivo borroso se realizó en una red neuronal artificial monocapa semejante a la red de Hopfield (29) pero con conexiones autorrecurrentes. Los pesos asociados con las conexiones coincidieron con los pesos asignados en el mapa; las entradas son los valores asignados a la modalidad elegida dentro de cada variable y las salidas son las soluciones obtenidas por la red. La implementación y uso de la red se hizo en tres etapas: aprendizaje, calibración y validación. La solución de la red se obtuvo por medio de un producto de matrices [conjunto de valores asig-

nados a las modalidades por ($\times$) conjunto de pesos asignados a las conexiones] en Microsoft Excel®. El vector de valores conseguidos se procesó dentro de la neurona a través de su función de activación tangente hiperbólica para obtener en la salida una serie de valores en escala continua (solo interesó el valor que expresaba el riesgo de mortalidad). Las múltiples iteraciones que requirió la red para dar el valor final se monitorizaron en forma gráfica hasta observar el *plateau* que indicó cuándo el resultado final se "estabilizaba".

Paso 5. La recolección de los datos de los pacientes se realizó en forma prospectiva en una base de datos relacional *ad hoc*.

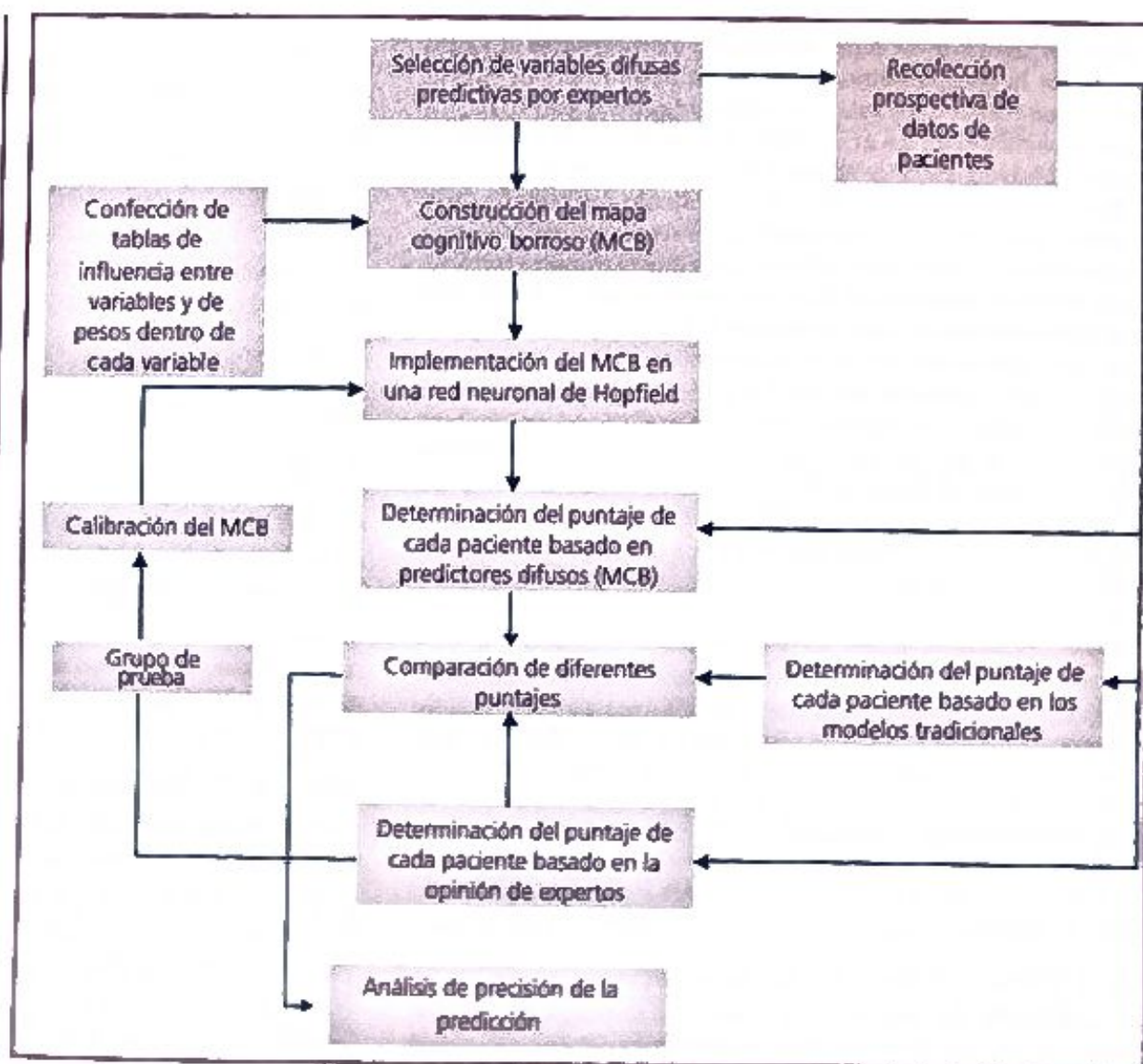
Paso 6. El experto evaluó a cada paciente en el preoperatorio y le asignó un riesgo de mortalidad en la escala ordinal bajo/moderado/alto en base al conjunto borroso de la Figura 2. Este se confeccionó teniendo en cuenta los estándares de calidad aceptados internacionalmente en el momento del estudio.

Paso 7. Determinación del puntaje o *score* de riesgo en un conjunto de prueba de pacientes usando los predictores borrosos. El mapa cognitivo borroso confeccionado en el Paso 4 se usó para determinar el riesgo quirúrgico de un conjunto de pacientes de prueba, a fin de calibrar sus resultados con los asignados por el experto. La calibración consistió en la modificación de los pesos de influencia entre variables e intra-variables adjudicados inicialmente en las tablas de influencia.

Paso 8. Determinación prospectiva del puntaje de riesgo del conjunto de validación usando el mapa cognitivo borroso ya calibrado. El modelo adjudicó a cada paciente un valor de riesgo en una escala continua y se transformó también a la misma escala ordinal bajo/moderado/alto.

Paso 9. Determinación del puntaje de riesgo de cada paciente de acuerdo con los modelos de Parsonnet, Ontario y EuroSCORE.

Fig. 1. Esquema del diseño del estudio.



Paso 10. Comparación de los resultados obtenidos con los distintos modelos. Se comparó el grado de concordancia para los conjuntos de prueba y validación entre "experto" y mapa cognitivo borroso. Posteriormente, dado que los puntajes originados a partir del mapa cognitivo borroso, Parsonnet, Ontario y EuroSCORE se hallan expresados en una escala continua, se realizó también un análisis de regresión y correlación lineal y no-lineal para relacionar el modelo borroso con los distintos puntajes.

Paso 11. Finalmente se evaluó el nivel de predicción alcanzada por el "experto" y su sistema borroso derivado y se comparó con el obtenido con los puntajes tradicionales.

Análisis estadístico

Para la comparación inicial de los resultados, entre la opinión del experto expresada en escala ordinal (grupos de riesgo bajo/moderado y alto) y los valores obtenidos con el mapa cognitivo borroso expresados en escala continua, se usó un análisis de varianza de un factor. Los límites puntuales entre los grupos con riesgo bajo/moderado y moderado/alto se optimizarán con el empleo de curvas ROC. Posteriormente, los valores de la escala continua del mapa cognitivo borroso se adjudicarán a cada conjunto borroso para realizar el correspondiente análisis de concordancia.

El análisis de concordancia para las tres categorías de riesgo, bajo/moderado/alto, se realizó con el coeficiente C corregido por agrupamiento; su significación se calculó con la prueba de chi cuadrado. Adicionalmente se calculó el índice de concordancia de kappa de Cohen ponderada y los correspondientes intervalos de confianza. La ponderación de kappa se realizó con el esquema de pesos cuadráticos. Para analizar el tipo y el grado de correlación entre los valores continuos obtenidos con los diferentes puntajes clásicos con respecto a la solución del mapa cognitivo borroso se empleó un análisis de regresión y correlación lineal y no-lineal. Finalmente se estimó la precisión de la predicción con los diferentes modelos mediante la comparación de las áreas obtenidas con curvas ROC y sus respectivos intervalos de confianza del 95% ($IC_{95\%}$). La comparación de las áreas ROC se realizó con el método de Hanley-McNeil. Para los cálculos se emplearon los paquetes de software Epidat 2.1* y SPSS 7.5*. El cálculo del tamaño de la muestra para el análisis de concordancia con kappa ponderada se realizó con la fórmula que considera que el número mínimo de individuos observados debe ser mayor que $2c^2$, donde c es la cantidad de categorías (celdas de la tabla de contingencia). Así, para una tabla 3×3 se tiene $2 \times 9^2 = 162$ individuos. El grupo de validación estuvo formado por una muestra del doble del tamaño calculado $162 \times 2 = 320$ pacientes para compensar la posible pérdida de datos. No existe consenso aún en cuanto al número de observaciones necesarias para calibrar o probar una red de Hopfield, por lo que se empleó una muestra de prueba de por lo menos un tercio del tamaño calculado originalmente.

Consideraciones éticas

El uso de los registros médicos para este estudio se realizó con la respectiva aprobación de los Comités Institucionales y de acuerdo con las normativas regulatorias de la ANMAT para estudios clínicos observacionales prospectivos (Disposición 5330/97). Asimismo, los pacientes dieron su consentimiento para el uso de esta información.

RESULTADOS

En la Tabla 1 se resumen las características basales de la población de acuerdo con sus variables preoperatorias e intraquirúrgicas, así como la morbilidad

Tabla 1. Características de la población. Variables preoperatorias, intraquirúrgicas y resultados operatorios inmediatos (n = 450)

Variabes	n (%)
Preoperatorio e Intraoperatorio:	
Edad, años (media $\pm$ DE)	63,7 $\pm$ 9,83
Sexo femenino	108 (24,0)
Diabetes	88 (19,6)
Hipertensión	284 (63,1)
Obesidad	85 (18,9)
Insuficiencia cardíaca	51 (11,3)
ACV	21 (4,7)
Enfermedad pulmonar	47 (10,4)
Disfunción renal	27 (6,0)
Anemia	12 (2,7)
Angor inestable	223 (49,6)
Infarto reciente	36 (8,0)
Reoperación	14 (3,1)
Disfunción VI moderada-grave	149 (33,1)
Riesgos calculados con: *	
Parsonnet	9,10 (8,29 a 10,5)
Ontario	4,39 (4,02 a 4,80)
EuroSCORE	6,32 (5,62 a 6,87)
Tipo de cirugía:	
Coronaria	315 (70,0)
Valvular	82 (18,2)
Combinada	56 (12,4)
Cirugía de urgencia	80 (17,8)
CRM sin CEC**	161 (51,1)
Posoperatorio:	
Mortalidad	21 (4,7)
Extubación en quirófano	374 (83,1)
Complicaciones:	
Reoperación por sangrado	8 (1,8)
Infarto (tipo Q)**	8 (2,5)
ACV	9 (2,0)
Diálisis	8 (1,8)
Mediastinitis	6 (1,3)

* Calculado sobre el conjunto de validación (n = 360).

** Calculado sobre las cirugías coronarias exclusivamente (n = 315).

Los riesgos calculados se presentan como media $\pm$ intervalo de confianza del 95%.

ACV: Accidente cerebrovascular. VI: Ventrículo izquierdo. DE: Desviación estándar. CRM: Cirugía de revascularización miocárdica. CEC: Circulación extracorpórea.

de los procedimientos. De acuerdo con la mortalidad global observada del 4,7%, se constata que los riesgos esperados pronosticados por los puntajes de Ontario y EuroSCORE son los más aproximados a este valor dato que luego fue corroborado con las áreas ROC. Por su parte, la distribución de los factores de riesgo preoperatorio y las complicaciones posquirúrgicas son similares a las de otras series publicadas.

Desarrollo del modelo

De acuerdo con las variables difusas seleccionadas por el experto y con las relaciones de influencia entre dichas variables para predecir mortalidad poscirugía cardíaca, se desarrolló el mapa cognitivo borroso (véase Figura en Material suplementario en la web). El mapa cognitivo borroso muestra los factores de riesgo preoperatorio y sus relaciones, que a juicio del "experto" pueden predecir la mortalidad hospitalaria de este tipo de cirugía. Los valores entre los nodos corresponden al peso asignado a cada conexión basada en las tablas de influencia. Los valores cercanos a 1 (uno) indican que dicha variable tiene mucha influencia para predecir mortalidad, mientras que una cifra próxima a 0 (cero) señala el poco peso de ese factor para pronosticar el evento adverso. Todas las variables tienen distinto grado de conexión y confluyen en el nodo central de salida, el cual va a predecir el riesgo de mortalidad de ese paciente en particular de acuerdo con la presencia o ausencia de los factores de riesgo señalados. Posteriormente, este mapa cognitivo borroso fue implementado en una red de Hopfield con conexiones autorrecurrentes. Para la etapa de aprendizaje y calibración del modelo se usó una muestra de 90 pacientes cuyos datos se incorporaron a la matriz de entrada de la red y, después de 12 iteraciones, sus resultados se expresaron con un puntaje de riesgo de mortalidad. El puntaje obtenido con la red se comparó con el riesgo pronosticado por el "experto" para cada paciente en base a los conjuntos borrosos para adjudicación del riesgo de mortalidad ya definidos en la Figura 2 (bajo/moderado/alto). En la Figura 3 se muestra esta comparación entre la opinión del "experto" expresada en escala ordinal *versus* los valores obtenidos con el mapa cognitivo borroso en escala continua. Las líneas transversales correspondientes a los puntajes 2,95 y 4,65 indican los puntos de corte para transformar la escala continua en ordinal (riesgos bajo, moderado y alto según el "experto"), y se obtuvieron mediante la optimización de dichos límites con curvas ROC (área ROC para el límite bajo/moderado: 0,89, IC_{95%} 0,809-0,978; área ROC para el límite moderado/alto: 0,87, IC_{95%} 0,788-0,955). Tras la determinación de estos límites, los riesgos calculados con el mapa cognitivo borroso en escala continua se transformaron a escala ordinal de acuerdo con los siguientes valores: riesgo bajo < 2,95, moderado 2,95 a 4,65 y alto > 4,65. Por último, para finalizar esta etapa de aprendizaje y calibración del modelo basado en el mapa cognitivo borroso, se analizó la correlación entre sus resultados y los del "experto". El valor de concordancia de kappa fue 0,672 (IC_{95%} 0,553-0,791) y el coeficiente C de contingencia corregido por agrupamiento fue 0,81 ($p < 0,0001$).

Validación del modelo

El modelo ya calibrado en la etapa de desarrollo se usó para predecir los resultados quirúrgicos de un conjunto de validación compuesto por 360 pacientes, a quienes se les determinó individualmente el puntaje asignado

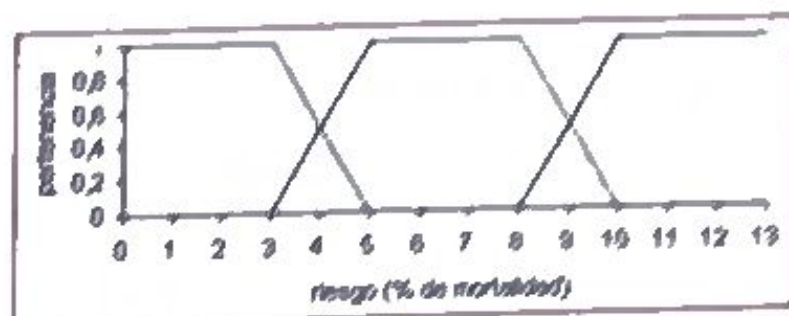


Fig. 2. Conjuntos borrosos usados para asignar riesgo por parte del "experto".

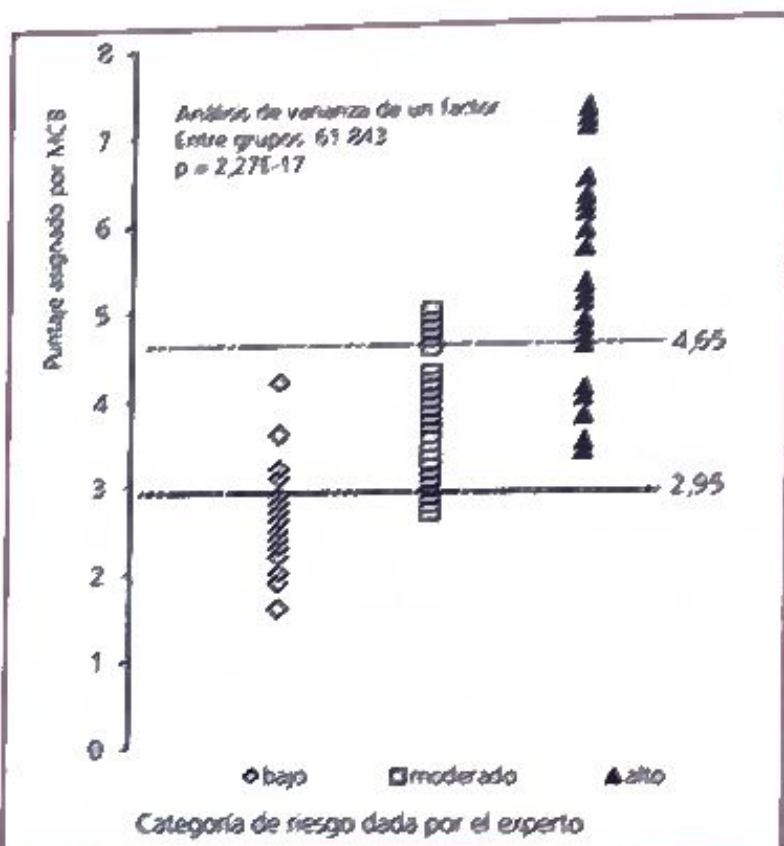


Fig. 3. Comparación inicial (conjunto de prueba o de aprendizaje y calibración) de la opinión del "experto" expresada en escala ordinal versus los valores obtenidos con el mapa cognitivo borroso (MCB) en escala continua. Las líneas transversales correspondientes a los puntajes 2,95 y 4,65 indican los puntos de corte para transformar la escala continua en ordinal (riesgos bajo, moderado y alto según el "experto"). Comparación estadística con análisis de varianza de un factor.

por el mapa cognitivo borroso. Adicionalmente se calcularon para cada paciente los riesgos pronosticados por tres puntajes de uso habitual en nuestro medio: Parsonnet, Ontario y EuroSCORE logístico. Posteriormente se calculó la precisión de cada modelo con curvas ROC y se comparó con el método de Hanley-McNeil según se muestra en la Figura 4. Los valores de las áreas ROC demuestran que el mapa cognitivo borroso tuvo por lo menos el mismo desempeño para predecir mortalidad hospitalaria poscirugía cardíaca en el conjunto de validación, aunque su área ROC fue la mejor. El área ROC del conjunto de validación de acuerdo con el desempeño del "experto" fue 0,767 (IC_{95%} 0,682-0,850). De acuerdo con el rango de puntaje obtenido con el mapa

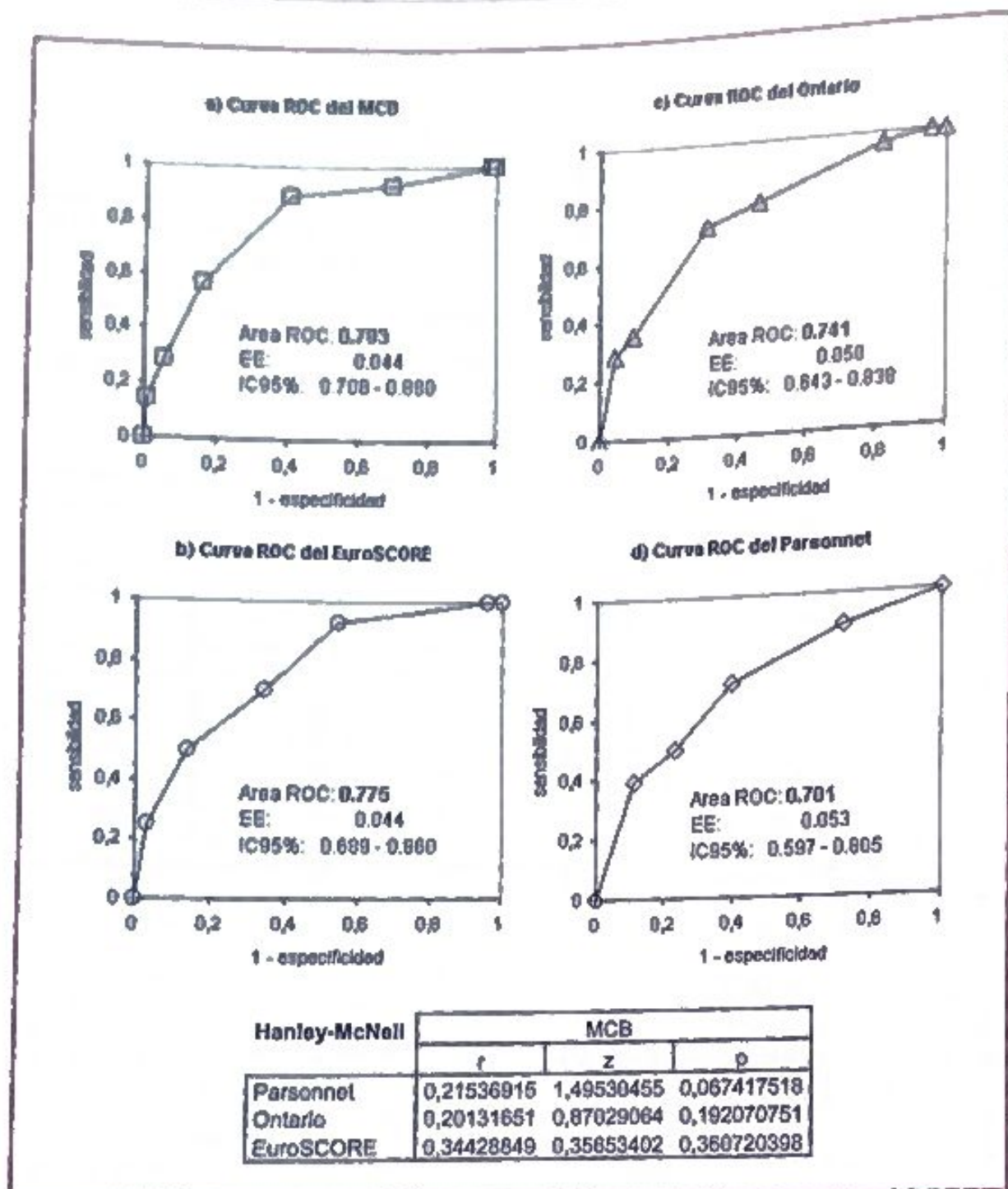


Fig. 4. Comparación de áreas ROC determinadas para el mapa cognitivo borroso (MCB) y los distintos puntajes de riesgo en el conjunto de validación. Tabla al pie con los valores de correlación (r) y significación (p) de acuerdo con el método de Hanley-McNeill.

cognitivo borroso se confeccionó la Figura 5, que sirve para determinar la mortalidad esperada de acuerdo con el valor asignado a cada paciente o grupo de pacientes.

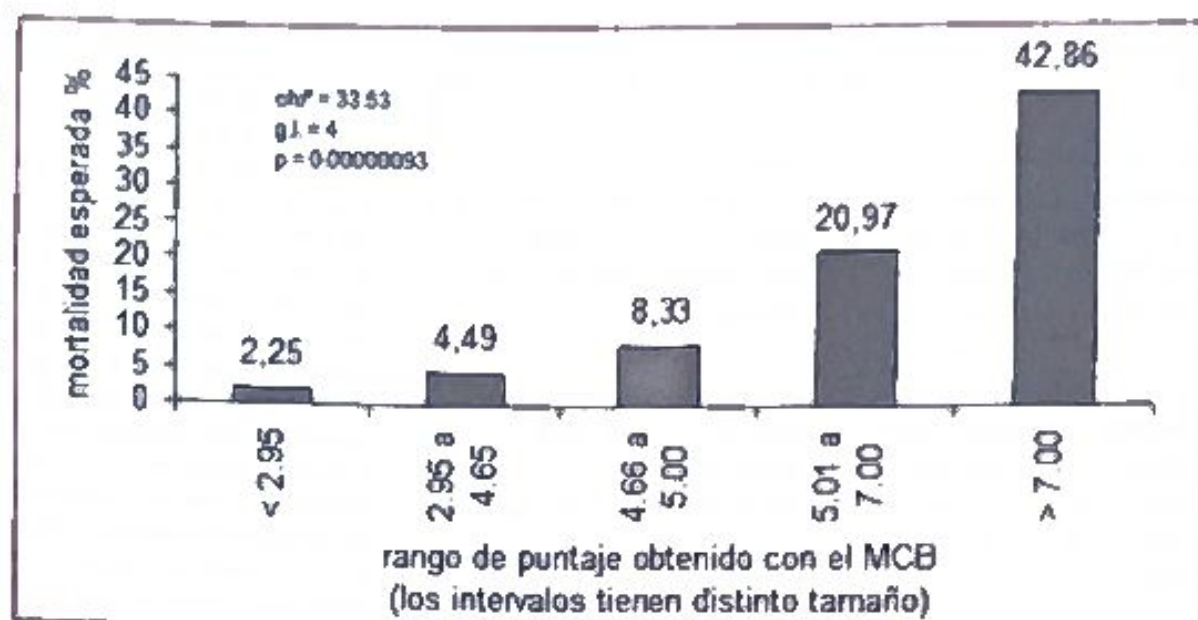
DISCUSIÓN

En este estudio fue posible desarrollar y validar un sistema de ajuste por riesgo basado en lógica borrosa para predecir la mortalidad hospitalaria de la cirugía cardíaca. Este modelo no solo imitó los resultados obtenidos por el "experto", sino que también, por lo menos, tuvo el mismo desempeño que otros puntajes de riesgo de uso habitual en nuestro medio al momento del estudio. Potencialmente, y dado que el modelo basado en lógica borrosa obtuvo comparativamente la mejor área ROC, sería posible conseguir aun un mejor desempeño que el resto de los puntajes mediante el ajuste de sus variables y relaciones, o aumentando el tamaño de la muestra.

El punto más destacable de la hipótesis de este trabajo es la característica incertidumbre con la que se maneja la obtención de los datos en la práctica médica. Las decisiones médicas suelen considerar un volumen elevado de información proveniente de distintas fuentes, a saber, los datos provistos por el paciente en la anamnesis, el examen físico, los valores de laboratorio y otros estudios complementarios, los que a su vez pueden ser incompletos, vagos, desconocidos o incluso contradictorios. Es evidente que, en este contexto, el proceso de decisión médica requiere herramientas que sean capaces de manejar la complejidad del problema y, a su vez, que puedan inferir un resultado u opinión.

La posibilidad de manejar el conocimiento experto como si se tratara de datos discretos o continuos analizables algebraicamente es la ventaja fundamental que aporta la lógica difusa y sus métodos de modelización, como son el mapa cognitivo borroso y la red neuronal artificial de Hopfield. El uso de la lógica difusa, que sir-

Fig. 5. Mortalidad esperada de acuerdo con el puntaje obtenido con el mapa cognitivo borroso (MCB).



ve de marco a los denominados modelos posibilísticos, es capaz de gestionar tanto la incertidumbre como la imprecisión en los sistemas expertos. En este caso, la base de conocimientos del sistema experto contiene una serie de reglas y hechos cargados con anterioridad, que deben generar una conclusión asociada con cierto grado de certeza o probabilidad de que suceda. Por su parte, el mapa cognitivo borroso se construye con una serie de nodos que representan los conceptos relevantes o variables del sistema, y unas conexiones entre ellos que muestran la relación causal o influencia entre las variables. Cada una de estas últimas está representada por un conjunto borroso a través del cual se puede integrar el conocimiento del experto. Presumiblemente, estos modelos siguen un planteo similar al razonamiento y proceso de decisión humano cuando trata de abordar el tratamiento de los denominados sistemas complejos.

Los resultados de este trabajo muestran que es posible usar el conocimiento del experto cuando este lo expresa en forma coloquial, manifestando la importancia de cada dato clínico como de poca, moderada o alta influencia sobre el resultado buscado. De esta forma, no se obliga al experto a adjudicar un valor numérico concreto de fuerza o probabilidad de asociación entre los hechos, sino que se le permite la transmisión del conocimiento en forma difusa. A partir de esta información pudo modelarse un sistema que compitió eficazmente con los otros modelos clásicos basados en regresión logística o análisis bayesiano. En resumen, los datos clínicos obtenidos de la observación directa del paciente ingresados en un formulario *ad hoc*, permitiría determinar la mortalidad esperada de acuerdo con el valor asignado a cada paciente o grupo de pacientes.

Existen algunos aspectos metodológicos que deben discutirse. Por un lado, el conocimiento del experto no se basa exclusivamente en su experiencia clínica, sino también en todo lo recolectado y aprendido de la bibliografía sobre el tema. Así que es muy probable que su conocimiento incluya, aunque sea parcialmente, la información sugerida por gran parte de la bibliografía

sobre modelos de ajuste por riesgo a la que él haya tenido acceso. De esta forma podría existir algún tipo de sesgo cuando se compara el desempeño del mapa cognitivo borroso con el de los otros puntajes. El segundo punto de importancia es la forma de adjudicación de los pesos a las tablas de influencias entre variables. La opinión del "experto" se manifiesta como un conjunto borroso; a partir de allí el peso adjudicado se deberá ajustar manualmente (calibración) hasta obtener la salida más estable de la red de Hopfield. Este proceso debe realizarse mediante "prueba y error", ya que no existe un método o algoritmo alternativo para resolverlo.

El modelo desarrollado en este trabajo constituye un verdadero sistema experto para la determinación del riesgo de la cirugía cardíaca, y además solo requiere ser alimentado con datos o información clínica extraídos a partir de la percepción del médico en su habitual contexto de incertidumbre. De ninguna manera pretende reemplazar el juicio clínico del médico, sino servir como sistema de apoyo para la toma de decisiones.

La mayoría de los trabajos de investigación en medicina relacionados con la aplicación de la lógica borrosa están habitualmente emparentados con el área de la inteligencia artificial. La lógica borrosa se ha usado para la descripción de la anatomía del sistema nervioso, (18) como sistema de apoyo para la toma de decisiones en cuidados paliativos, (19) para predecir el riesgo de sangrado después de una amigdalectomía, (20) para monitorizar la ventilación mecánica en niños, (30) en el análisis del temblor de origen neurológico, (31) para el procesamiento de imágenes, (30, 32) en investigación cardiovascular y caracterización de tipos de accidentes cerebrovasculares, (33) para la monitorización y control en anestesia, (34) etc. Por su parte, las publicaciones con aplicación de mapas cognitivos borrosos incluyen modelos para el tratamiento de la diabetes, (35) diagnóstico diferencial de alteraciones urinarias y de dificultades de aprendizaje, (36, 37) sistemas expertos en dolor abdominal agudo (38) y diagnóstico médico. (39)

Limitaciones

La validez temporal y posiblemente regional del uso de un puntaje o score impone la reevaluación constante del sistema de ajuste por riesgo. Así, un puntaje adecuado para evaluar un grupo de pacientes quirúrgicos en un momento dado podría exagerar o subvalorar en el futuro el riesgo esperado de otro grupo, cuando los estándares de calidad del momento exijan mejores resultados. Es por eso que la modelización para predecir el riesgo debe ser un proceso iterativo en el tiempo para adaptar el sistema a los nuevos niveles de exigencia y calidad. Otra limitación es que el modelo basado en un mapa cognitivo borroso trabaja como una "caja negra", o sea que su funcionamiento no es transparente al usuario y por ende no es posible seguir la evolución del proceso hasta la obtención del resultado; es decir que el sistema no explica o justifica su opinión. Por último, la modelización que permiten estos sistemas se basan en el conocimiento experto, que también puede ser incorrecto o por lo menos de calidad inferior al conocimiento obtenido por otros métodos, como el análisis estadístico multivariante.

CONCLUSIONES

La posibilidad de contar con una metodología y un sistema capaz de aprovechar este conocimiento experto fue investigada en este trabajo mediante el uso de la lógica borrosa y la implementación de un sistema experto. En este estudio fue posible desarrollar y validar un sistema de ajuste por riesgo basado en lógica borrosa para predecir la mortalidad hospitalaria de la cirugía cardíaca. El modelo no solo imitó los resultados obtenidos por el "experto", sino que también, por lo menos, tuvo el mismo desempeño que otros puntajes de riesgo. En una escala mayor, la lógica difusa y los mapas cognitivos borrosos podrían ser útiles para desarrollar sistemas distribuidos capaces de combinar información de distinto origen. Por último, los modelos basados en lógica borrosa podrían formar parte de sistemas clasificadores múltiples que también incluyan el abordaje estadístico tradicional, algoritmos diagnósticos y árboles de decisión que en conjunto podrían mejorar la precisión diagnóstica del modelo. (40)

Declaración de conflicto de intereses

Los autores declaran que no poseen conflicto de intereses.

(Véanse formularios de conflicto de intereses de los autores en la web/ Material suplementario).

BIBLIOGRAFÍA

1. Cowan JA, Dimick JB, Wainess RM, Upchurch GR Jr, Thompson BG. Outcomes after cerebral aneurysm clip occlusion in the United States: the need for evidence-based hospital referral. *J Neurosurg* 2003;99:947-52. <http://doi.org/fmn7hg>
2. Dimick JB, Cowan JA Jr, Allawadi G, Wainess RM, Upchurch GR Jr. National variation in operative mortality rates for esophageal resection and the need for quality improvement. *Arch Surg* 2003;138:1305-9. <http://doi.org/d93cgs>

3. Cowan JA, Dimick JB, Henke PK, Huber TS, Stanley JC, Upchurch GR Jr. Surgical treatment of intact thoracoabdominal aortic aneurysms in the United States: hospital and surgeon volume-related outcomes. *J Vasc Surg* 2003;37:1169-74. <http://doi.org/dtkh6>
4. Dimick JB, Pronovost PJ, Cowan JA, Lipsett PA. Complications and costs after high-risk surgery: where should we focus quality improvement initiatives? *J Am Coll Surg* 2003;196:671-8. <http://doi.org/dro4a3>
5. Nashef SA, Roques F, Sharples LD, Nilsson J, Smith C, Goldstone AR, Lockowandt U. EuroSCORE II. *Eur J Cardiothorac Surg* 2012;41:734-44; discussion 744-5. <http://doi.org/pvr>
6. Lee JH, Abdelhady K, Capdeville M. Clinical outcomes and resource usage in 100 consecutive patients after off-pump coronary bypass procedures. *Surgery* 2000;128:548-55. <http://doi.org/bd5m5x>
7. Parsonnet V, Dean D, Bernstein AD. A method of uniform stratification of risk for evaluating the results of surgery in acquired adult heart disease. *Circulation* 1989;79(Suppl I):I-3-I-12.
8. Tu JV, Jaglal SB, Naylor CD and the Steering Committee of the Provincial Adult Cardiac Care Network of Ontario. Multicenter validation of a risk index for mortality, intensive care unit stay, and overall hospital length of stay after cardiac surgery. *Circulation* 1995;91:677-84. <http://doi.org/59f>
9. Nashef SAM, Roques F, Michel P, Gauducheau E, Lemesbow S, Salamon R, the EuroSCORE study group. European system for cardiac operative risk evaluation (EuroSCORE). *Eur J Cardiothorac Surg* 1999;16:9-13. <http://doi.org/b9r62k>
10. O'Connor GT, Plume SK, Olmstead EM, Coffin LH, Morton JR, Maloney CT, et al. Multivariate prediction of in-hospital mortality associated with coronary artery bypass graft surgery. *Circulation* 1992;85:2110-8. <http://doi.org/59h>
11. Higgins TL, Estafanous FG, Loop FD, Beck GJ, Blum JM, Peranandi L. Stratification of morbidity and mortality outcome by preoperative risk factors in coronary artery bypass patients. A clinical severity score. *JAMA* 1992;267:2344-8. <http://doi.org/b8j9bd>
12. Grover FL, Hammermeister KE, Burchfiel C, and Cardiac Surgeons of the Department of Veterans Affairs. Initial report of the Veterans Administration preoperative risk assessment study for cardiac surgery. *Ann Thorac Surg* 1990;50:12-28. <http://doi.org/bj8h92>
13. Keogh BE. Cardiac surgical mortality: the tip of the quality assurance iceberg. *Heart* 2000;84:7-8. <http://doi.org/ckm3qx>
14. Pons JMV. El ajuste de riesgo en la medida de los resultados de la cirugía cardíaca. *Cir Cardiov* 1997;4:4-14.
15. Combettes PL. The foundations of set theoretic estimation. *Proc IEEE* 1993;81:2. <http://doi.org/brkcnx>
16. Zadeh LA. Fuzzy sets. *Information and Control* 1965;8:338-53. <http://doi.org/fef6rc>
17. Sadgh-Zadeh K. The fuzzy revolution: goodbye to the Aristotelian Weltanschauung. *Artif Intell Med* 2001;21:1-25. <http://doi.org/dmr699>
18. Axer H, Sudfeld D, van Keyserlingk DG, Berks G. Fuzzy sets in human anatomy. *Artif Intell Med* 2001;21:147-52. <http://doi.org/c6623v>
19. Beliakov G, Warren J. Fuzzy logic for decision support in chronic care. *Artif Intell Med* 2001;21:209-13. <http://doi.org/ctf2gx>
20. Pizzi NJ. Bleeding predisposition assessments in tonsillectomy/adenoidectomy patients using fuzzy interquartile encoded neural networks. *Artif Intell Med* 2001;21:65-90. <http://doi.org/b9vhr>
21. Mago VK, Mehta R, Woolrych R, Papageorgiou EI. Supporting meningitis diagnosis amongst infants and children through the use of fuzzy cognitive mapping. *BMC Med Inform Decis Mak* 2012;12:98. <http://doi.org/59j>
22. Turna A, Mercan CA, Bedirhan MA. Prediction of morbidity after lung resection in patients with lung cancer using fuzzy logic. *Thorac Cardiovasc Surg* 2005;53:368-74. <http://doi.org/dvnx7n>
23. Catto JW, Abbod MF, Connectionens DA, Larré S, Rosario DA, Hamdy FC. Neurofuzzy modeling to determine recurrence risk following radical cystectomy for nonmetastatic urothelial carcinoma of the bladder. *Clin Cancer Res* 2009;15:3150-5. <http://doi.org/bdx435>
24. Shamim MS, Enam SA, Qidwai U. Fuzzy Logic in neurosurgery predicting poor outcomes after lumbar disk surgery in 501 consecutive

- tive patients. *Surg Neurol* 2009;72:565-72; discussion 572. <http://doi.org/bh4m36>
25. Godil SS, Shamim MS, Enam SA, Qidwai U, Qadeer M, Sobani ZA. Cranial reconstruction after decompressive craniectomy: prediction of complications using fuzzy logic. *J Craniofac Surg* 2011;22:1307-11. <http://doi.org/cbxxrk>
26. Gutiérrez Moya E, González Camacho MC, Salmerón Silvera JL. Analyzing the impact of decisions in the scope of long term care by fuzzy cognitive maps, Spain. *Rev Esp Salud Pública* 2012;86:565-74.
27. Mago VK, Mehta R, Woolrych R, Papageorgiou EI. Supporting meningitis diagnosis amongst infants and children through the use of fuzzy cognitive mapping. *BMC Med Inform Decis Mak* 2012;12:98. <http://doi.org/59j>
28. Kosko B. Fuzzy Cognitive Maps. *International Journal of Man-Machine Studies* 1986;January:65-75. <http://doi.org/cpk5rz>
29. Hopfield J. Neural networks and physical systems with emergent collective computational abilities. *Proc Nat Acad Sc* 1982;79:2554-8. <http://doi.org/c2b78f>
30. Teodorescu HNL, Kandel A, Hall LO. Report of research activities in fuzzy AI and medicine at USF CSE. *Artif Intell Med* 2001;21:177-83. <http://doi.org/fkmw9f>
31. Teodorescu HNL, Kandel A. Non-linear analysis of tremor and applications. *Jpn J Biomed Eng* 1999;13:11-20.
32. Lu Y, Jiang T, Zang Y. Region growing method for analysis of functional MRI data. *NeuroImage* 2003;20:455-65. <http://doi.org/ft6hv3>
33. Helgason CM, Malik DS, Cheng SC, Jobe TH, Mordeson JN. Statistical versus fuzzy measures of variable interaction in patients with stroke. *Neuroepidemiology* 2001;20:77-84. <http://doi.org/bw9dwv>
34. Martin JF. Fuzzy control in anaesthesia. *J Clin Monitoring* 1994;10:77-80. <http://doi.org/derfnd>
35. Giles BG, Findlay CS, Haas G, Lafrance B, Laughing W, Pembleton S. Integrating conventional science and aboriginal perspectives on diabetes using fuzzy cognitive maps. *Soc Sci Med* 2007;64:562-76. <http://doi.org/b6mv9j>
36. de Moraes Lopes MH, Marin HF, Ortega NR, Massad E. Fuzzy logic model based on the differential nursing diagnosis of alterations in urinary elimination. *Stud Health Technol Inform* 2006;122:117-20.
37. Georgopoulos VC, Malandraki GA, Stylios CD. A fuzzy cognitive map approach to differential diagnosis of specific language impairment. *Artif Intell Med* 2003;29:261-78. <http://doi.org/fwn988>
38. Fathi-Torbaghan M, Meyer D. MEDUSA: a fuzzy expert system for medical diagnosis of acute abdominal pain. *Methods Inf Med* 1994;33:522-9.
39. John RI, Innocent PR. Modeling uncertainty in clinical diagnosis using fuzzy logic. *IEEE Trans Syst Man Cybern B Cybern* 2005;35:1340-50. <http://doi.org/dkkff4>
40. Sboner A, Eccher C, Blanzieri E, Bauer P, Cristofolini M, Zumiani G, et al. A multiple classifier system for early melanoma diagnosis. *Artif Intell Med* 2003;27:29-44. <http://doi.org/bjkvs6>