



Diagnóstico de riesgo de aterogénesis asistido por lógica borrosa

O.A. Rosas-Jaimes¹, A.L. Alonso-García², J. Mas-Oliva² y L. Alvarez-Icaza¹

¹Instituto de Ingeniería, UNAM e ²Instituto de Fisiología Celular, UNAM

E-mails: ORosasJ@iingen.unam.mx, aalonso@ifisiol.unam.mx, jmas@ifisiol.unam.mx y alvar@pumas.iingen.unam.mx

(Recibido: octubre de 2004; aceptado: mayo de 2006)

Resumen

Se describe un sistema de inferencias borrosas que asiste en el diagnóstico de riesgo de aterogénesis, a partir del análisis de muestras de plasma humano. El sistema utiliza por primera vez la proteína transferidora de ésteres de colesterol (CETP), además de las concentraciones en plasma de colesterol total, lipoproteínas de baja densidad y el índice aterogénico para proponer un método de diagnóstico no intrusivo que auxilia en la detección temprana de riesgo de aterogénesis a un costo razonable.

Descriptores: Diagnóstico clínico asistido por computadora, CETP, LDL, IA, colesterol, lógica borrosa, sistemas expertos.

Abstract

A fuzzy system based on human plasma analysis to assist in the diagnosis of atherogenesis risk is described. The system introduces a new factor, the cholesterol ester transfer protein, that together with concentrations of total cholesterol, low density lipoproteins and the atherogenic index, allows to propose a non-intrusive diagnostic method that is very helpful in low cost early detection of atherogenesis risk.

Keywords: Computer aided clinical diagnosis, CETP, LDL, IA, cholesterol, fuzzy logic, expert systems.

Introducción

La lógica borrosa y la medicina tienen en común el manejo de información con un cierto nivel de incertidumbre. Si bien es cierto que el profesional médico debe basar una gran parte de sus diagnósticos en resultados de análisis cuantitativos, también lo es el hecho de que parte de sus diagnósticos están basados en decisiones que tienen que ver con la experiencia e intuición propias del experto, características cuya naturaleza es subjetiva.

Otro punto de convergencia entre la medicina y la lógica borrosa, es la forma de establecer un razonamiento basado en conocimiento. En ambos casos, a través de datos –los síntomas que un

paciente muestra o los resultados de análisis– y de un conjunto de reglas relacionadas con ese conocimiento, se busca establecer un diagnóstico. A partir de este último, se prescribe un tratamiento que permita corregir desbalances y conseguir que un paciente vuelva a una situación de equilibrio deseable. En la mayoría de los casos, tanto en la medicina como en la lógica borrosa, esto se logra sin un conocimiento total y exacto del funcionamiento de los diferentes sistemas bajo análisis.

La incursión de la lógica borrosa en la medicina surge poco después del establecimiento de aquella como campo de conocimiento, para constituirse, junto con las redes neurales, los algoritmos genéticos, los sistemas bayesianos, etcétera, así como híbridos de todos ellos, en herramientas que

conformarían los primeros sistemas expertos de tipo médico (Hudson y Cohen, 1994). Estos sistemas comenzaron por asistir al profesional en la valoración de síntomas exhibidos por el paciente (Tseng y Teo, 1994; Nguyen *et al.*, 2001). Posteriormente, la lógica borrosa, en unión con otras técnicas, se ha extendido al análisis y reconocimiento de patrones en imágenes utilizadas en diagnóstico, tema en donde resulta extensa la literatura (Udupa *et al.*, 1997, Choi y Krishnapuramm, 1995, por citar algunos ejemplos). Sin embargo, el número de trabajos en donde se utilizan conceptos borrosos para asistir al diagnóstico con base en análisis clínicos es más reducido (Gorzalczany y McLeish, 1992).

En este trabajo se presenta un sistema borroso basado en reglas para determinar el estado de riesgo de aterogénesis, o desarrollo de aterosclerosis, en sujetos a los que se les ha practicado análisis clínico sanguíneo. Existe un trabajo con objetivo similar (Sidaoui y Pacheco, 1999), que realiza un reconocimiento de patrones en imágenes obtenidas mediante catéteres introducidos en arterias. Aún cuando el uso de la metodología propuesta en dicho artículo produce buenos resultados en el diagnóstico, dada la naturaleza invasiva del procedimiento, sólo resulta útil en casos en donde el daño ya es irreversible.

El método propuesto, plantea el uso de análisis de plasma como base para elaborar el diagnóstico, ya que este tipo de análisis es de uso más frecuente por su relativo bajo costo y su carácter no invasivo. Estos análisis pueden proporcionar información acerca de los compuestos que circulan en el torrente sanguíneo, con la que es posible obtener diagnósticos tempranos.

El sistema borroso aquí expuesto utiliza por primera vez los resultados de una investigación realizada en el Instituto de Fisiología Celular (IFC) de la UNAM, en donde se ha desarrollado y validado un método de diagnóstico de riesgo de aterogénesis con base en la cantidad de Proteína Transferidora de Ésteres de Colesterol (*CETP*) presente en plasma. El propósito de dicha investigación es facilitar la detección temprana de sujetos con riesgo de aterogénesis, además de reducir tanto el tiempo como el costo de dicho

diagnóstico. Este método ha sido registrado y su patente se encuentra en proceso en las oficinas Mexicana, Estadounidense, Canadiense y de la Comunidad Europea (Alonso y Mas-Oliva, 2002).

Un sistema basado en la inclusión de la *CETP* para la formulación de diagnósticos tempranos de riesgo de aterogénesis no ha sido reportado anteriormente. Es importante mencionar que los resultados reportados en este artículo están basados en una amplia investigación clínica. El sistema se desarrolló con dos propósitos fundamentales:

- 1) Corroborar o modificar, con una metodología alternativa, los resultados obtenidos por medios convencionales de inferencia y

- 2) Proporcionar a los investigadores y médicos usuarios de la metodología una herramienta de fácil uso para poder calibrar cambios o realizar adiciones en un futuro a este sistema diagnóstico o alguno similar.

Este trabajo no es una aportación a la teoría de sistemas difusos. Se trata de un trabajo multidisciplinario que intenta combinar metodologías de disciplinas distintas para desarrollar un sistema de diagnóstico que apoye a los profesionales de la medicina que desde la investigación o práctica clínica se enfrentan a los padecimientos de aterogénesis en números crecientes.

Para la exposición de este trabajo, la Sección 2 trata del método de diagnóstico desarrollado en el IFC y describe las variables que éste involucra; el diagnóstico se transforma en el sistema de inferencias borrosas de que trata este artículo y cuyo diseño se aborda en la Sección 3. En la Sección 4 se muestran algunos resultados obtenidos de sujetos a los que se les ha practicado análisis sanguíneo. Por último, se presentan conclusiones surgidas de este trabajo.

Generalidades del diagnóstico de riesgo de aterogénesis

La aterogénesis es un proceso de formación de placas, principalmente de lípidos, en las paredes arteriales, que conduce al estrechamiento del

diámetro interior o luz de estos vasos sanguíneos. Las consecuencias que este fenómeno produce van desde el aumento de la tensión arterial hasta el desarrollo de condiciones que ocasionan gangrena, embolia o infarto del miocardio, padecimientos ocasionados por la obstrucción del paso de sangre hacia miembros inferiores, cerebro o corazón y que muchas veces provocan la muerte.

Actualmente, la manera más común de establecer un diagnóstico médico de esta enfermedad es mediante lineamientos que dependen, tanto de los niveles de lípidos como de la presencia o ausencia de enfermedad cardiovascular (ECV) establecida, así como de otros factores de riesgo coronario. Dichos factores indican la exposición del individuo a circunstancias que pueden favorecer el riesgo de aterogénesis. Para establecer si existe exposición del paciente a factores de riesgo de enfermedad coronaria, se realizan además, interrogatorios dirigidos, exploración, determinación del perfil de lípidos, electrocardiograma y radiografía de tórax (Alonso, 2003). Toda esta información completa un conjunto de datos que el médico debe tomar en cuenta y analizar para establecer diagnósticos y prescripciones.

Para el análisis de lípidos sanguíneos, existe un criterio ampliamente aceptado que mide las concentraciones en plasma de Colesterol Total (CT), de Lipoproteínas de Baja Densidad (LDL) y el Índice Aterogénico (IA), siendo éste calculado mediante el cociente $IA = HDL/CT$. En la tabla 1

se enlistan los rangos de estas variables relacionados con el diagnóstico de riesgo de aterogénesis.

Alonso *et al.* (2003) han encontrado que, además de estas variables, la Proteína Transferidora de Ésteres de Colesterol (CETP) juega un papel muy importante en la precisión y facilidad de establecimiento del diagnóstico de aterogénesis y afirman que debe ser incluida para llevarlo a cabo. Los autores citan cuatro categorías para el nivel de CETP plasmática, las cuales se muestran en la tabla 2.

En dicha tabla, la categoría *sobreexpresión* agrupa valores relacionados con CT y LDL, cuyas cantidades, más allá de valores considerados de *Alto Riesgo*, vuelven a caer en clasificaciones *deseable* o *límitrofe*. Alonso (2003), relaciona este hecho con la posibilidad de una deficiencia en la actividad de CETP, que estimula la sobreexpresión de dicha proteína. Es importante recalcar que las tablas 1 y 2 se basan en estudios de análisis clínicos para una población de pacientes mexicanos.

Conociendo las concentraciones de cada una de estas cuatro variables, es posible establecer en lo químico un diagnóstico del estado de riesgo de aterogénesis, el cual, combinado con otro tipo de datos como los arriba mencionados, permite al profesional establecer recomendaciones al paciente.

Generalmente, el médico decide si el paciente se encuentra en un riesgo calificado como *nulo*,

Tabla 1. Intervalos clínicos de CT, LDL e IA para el diagnóstico de aterogénesis

	Deseable	Límitrofe	Alto Riesgo
$CT \left[\frac{mg}{dl} \right]$	< 200	200 – 239	> 240
$LDL \left[\frac{mg}{dl} \right]$	< 130	130 – 160	> 160
IA	≤ 4.2		> 4.2

Tabla 2: Categorías de riesgo de aterogénesis para CETP $\left[\frac{\mu g}{ml} \right]$

	Deseable	Límitrofe	Alto Riesgo	Sobreexpresión
CETP	< 0.8	0.8 – 1.8	1.8 – 6.0	> 6.0

bajo o alto, a través de su conocimiento, el cual puede ser esquematizado mediante reglas de la forma:

$$\begin{aligned} & \text{si } CT \text{ es } CT_{deseable} \text{ y} \\ & LDL \text{ es } LDL_{deseable} \text{ y} \\ & IA \text{ es } IA_{deseable} \text{ y} \\ & CETP \text{ es } CETP_{deseable} \\ & \text{entonces } RIESGO \text{ es Sin Riesgo} \end{aligned} \quad (1)$$

Puede observarse que las cuatro variables clínicas están incluidas en una parte *antecedente* de la regla (1) y están relacionadas con una única variable de interés, *RIESGO*, que forma la parte consecuente de dicha regla. Esta última variable es el diagnóstico obtenido del procesamiento de las cantidades involucradas en la parte antecedente, a través del conocimiento del experto médico.

Diseño del sistema de inferencias borrosas

Para emular el conocimiento experto expuesto en la sección anterior, se *difunden*¹ las variables que formarán parte del antecedente de las reglas de inferencia: *CT*, *LDL*, *CETP* e *IA*. Por otro lado, la parte consecuente constará en este diseño de una sola variable, riesgo (*R*), a la cual se asignará un valor entre 0 y 100 para calificar el estado de riesgo de desarrollar aterogénesis en un paciente cualquiera dados sus *CT*, *LDL*, *CETP* e *IA*.

Para cada una de estas variables se definen, a continuación, las siguientes funciones de pertenencia a los conjuntos borrosos:

$$\mu_{CT_{deseable}}(CT) = \begin{cases} 1 & \text{si } CT < 200 \\ \frac{240-CT}{240-200} & \text{si } 200 \leq CT \leq 240 \\ 0 & \text{si } CT > 240 \end{cases} \quad (2a)$$

$$\mu_{CT_{at\ riesgo}}(CT) = \begin{cases} 0 & \text{si } CT < 200 \\ \frac{CT-200}{240-200} & \text{si } 200 \leq CT \leq 240 \\ 1 & \text{si } CT > 240 \end{cases} \quad (2b)$$

Para $75 \leq CT \leq 600$

$$\mu_{LDL_{deseable}}(LDL) = \begin{cases} 1 & \text{si } LDL < 130 \\ \frac{160-LDL}{160-130} & \text{si } 130 \leq LDL \leq 160 \\ 0 & \text{si } LDL > 160 \end{cases} \quad (3a)$$

$$\mu_{LDL_{at\ riesgo}}(LDL) = \begin{cases} 0 & \text{si } LDL < 130 \\ \frac{LDL-130}{160-130} & \text{si } 130 \leq LDL \leq 160 \\ 1 & \text{si } LDL > 160 \end{cases} \quad (3b)$$

Para $50 \leq LDL \leq 250$

$$\mu_{CETP_{deseable}}(CETP) = \begin{cases} 1 & \text{si } CETP < 0.8 \\ \frac{1.8-CETP}{1.8-0.8} & \text{si } 0.8 \leq CETP \leq 1.8 \\ 0 & \text{si } CETP > 1.8 \end{cases} \quad (4a)$$

$$\mu_{CETP_{at\ riesgo}}(CETP) = \begin{cases} 0 & \text{si } CETP < 0.8 \\ \frac{CETP-0.8}{1.8-0.8} & \text{si } 0.8 \leq CETP \leq 1.8 \\ 1 & \text{si } 1.8 \leq CETP \leq 6.0 \\ \frac{6.5-CETP}{6.5-6.0} & \text{si } 6.0 \leq CETP \leq 6.5 \\ 0 & \text{si } CETP > 6.5 \end{cases} \quad (4b)$$

$$\mu_{CETP_{Ps.exp.}}(CETP) = \begin{cases} 0 & \text{si } CETP < 6.0 \\ \frac{CETP-6.0}{6.5-6.0} & \text{si } 6.0 \leq CETP \leq 6.5 \\ 1 & \text{si } CETP > 6.5 \end{cases} \quad (4c)$$

Para $0.0 \leq CETP \leq 10.0$

$$\mu_{IA_{deseable}}(IA) = \frac{8.4-IA}{8.4-0} \quad \text{si } IA < 8.4 \quad (5a)$$

$$\mu_{IA_{at\ riesgo}}(IA) = \begin{cases} \frac{IA-0}{8.4-0} & \text{si } IA \leq 8.4 \\ 1 & \text{si } IA > 8.4 \end{cases} \quad (5b)$$

Para $0.0 \leq IA \leq 16.0$

$$\mu_{No\ Obse\ ri\ v\ ad}\ (R) = 1 \quad \text{si } R = -2 \quad (6a)$$

$$\mu_{S\ in\ riesgo}\ (R) = \begin{cases} \frac{20-R}{20-0} & \text{si } R \leq 20 \\ 0 & \text{si } R > 20 \end{cases} \quad (6b)$$

¹ *Difundir* denota la conversión de una cantidad real en una borrosa.

$$\mu_{\text{BajoRiesgo}}(R) = \begin{cases} 0 & \text{si } R < 20 \\ \frac{R-20}{35-20} & \text{si } 20 \leq R \leq 35 \\ \frac{50-R}{50-35} & \text{si } 35 \leq R \leq 50 \\ 0 & \text{si } R > 50 \end{cases} \quad (6c)$$

$$\mu_{\text{AltoRiesgo}}(R) = \begin{cases} 0 & \text{si } R < 50 \\ \frac{R-50}{100-50} & \text{si } 50 \leq R \leq 100 \end{cases} \quad (6d)$$

Para $0 \leq R \leq 100$

Los valores $\mu_{\text{LDLdecahle}}(LDL)$, o $\mu_{\text{Alto riesgo}}(IA)$, por ejemplo, indican un valor de pertenencia a los conjuntos borrosos asociados a las variables lingüísticas *LDL* e *IA*, en este caso. La forma de estas funciones de pertenencia se muestra en la

figura 1. Nótese que los nombres asociados a ellas sugieren lingüísticamente los conceptos médicos asociados a dichas variables.

Por otra parte, las funciones de pertenencia en que se encuentra dividido el consecuente *R* dan a su vez, idea de la forma de calificar el riesgo de desarrollar aterosclerosis (Figura 2). Debido a que ciertas combinaciones de las variables antecedentes corresponden a situaciones contradictorias en la vida real (no es posible, por ejemplo, que un sujeto exhiba al mismo tiempo un *CT* muy alto junto a cifras de *LDL* muy bajas) fue necesario definir el valor borroso *NoObservado*, el cual dirige al riesgo a un valor de $R = -2$ para indicar que tales situaciones no corresponden con lo observado en clínica.

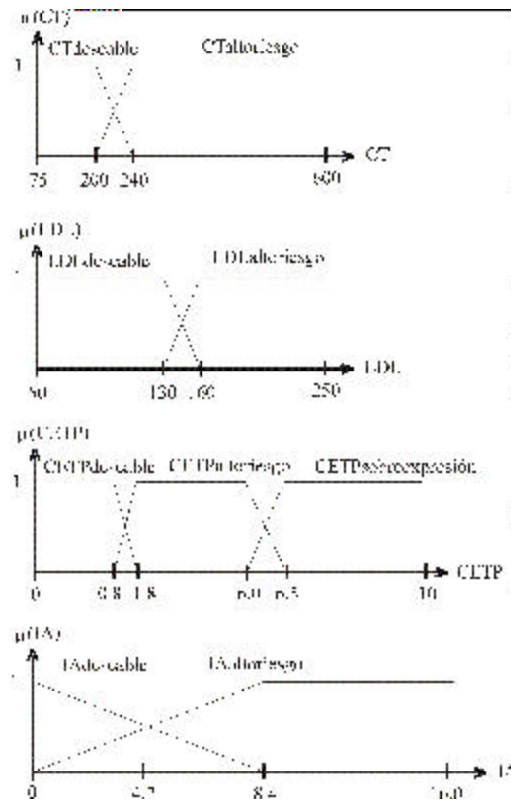


Figura 1. Funciones de pertenencia definidas para *CT*, *LDL*, *CETP* e *IA*

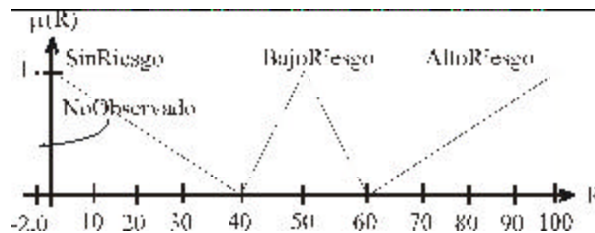


Figura 2. Funciones de pertenencia definidas para el estado de riesgo *R* de aterosclerosis

Los rangos para estas funciones de pertenencia se determinaron con base en los intervalos mostrados en la tabla 1, que es congruente con la forma en que la comunidad médica establece actualmente el diagnóstico de riesgo de aterogénesis, y se complementaron con los de la tabla 2, que resume los hallazgos analíticos sobre el papel de la CETP para el riesgo de aterogénesis. La base de razonamiento se establece a partir de un conjunto de reglas de inferencia que sirven de unión entre las variables antecedentes (*CT*, *LDL*, *CETP*, *IA*) y la consecuente (*R*). La forma triangular de las funciones de pertenencia no es determinante para obtener los resultados que se presentan más adelante; sin embargo, sí es de fácil entendimiento para la comunidad médica que potencialmente se ayudaría con este sistema. Es posible utilizar otras formas para definir la pertenencia a los conjuntos borrosos que, una vez calibradas, conducirán a resultados similares a los mostrados más adelante.

Como ejemplo de la forma en que funcionan los rangos de las funciones de pertenencia, considere las figuras 3 y 4 que muestran en el plano *CT* versus *CETP* los niveles de *LDL* e *IA*, respectivamente, los cuales resultaron de los análisis clínicos realizados a una muestra de 162 personas seleccionadas al azar. El tipo de marca indica el riesgo de aterogénesis determinado por *LDL*. Las líneas horizontales y verticales, representan los rangos

que se obtienen de las tablas 1 y 2. En estas figuras es posible visualizar fácilmente regiones que pueden ser calificadas de *Nulo*, *Bajo* o *Alto Riesgo*. Estas figuras, junto con el razonamiento experto, sirvieron como base para establecer reglas de inferencia con las cuales se relacionan las partes antecedente y consecuente.

Las 24 reglas de inferencia establecidas se muestran en el Apéndice. Las operaciones que se efectúan en cada una de estas reglas para obtener los valores de pertenencia de los consecuentes son de la forma

$$\mu_m(CT, LDL, CETP, IA, R) = \min\{\mu_i(CT) \wedge \mu_j(LDL) \wedge \mu_k(CETP) \wedge \mu_h(IA), \mu_p(R)\} \quad (7)$$

en donde los símbolos \wedge indican la operación "y" o "intersección"² el subíndice *i* define a alguna de las posibles funciones de pertenencia, por ejemplo, $\mu_{deseable}(CT)$ o $\mu_{altorriesgo}(CT)$ en las que se ha dividido el universo de discurso de *CT*. En forma equivalente, los subíndices *j*, *k*, *h* y *p* se relacionan con la funciones de pertenencia en que se parten los universos de discurso de las variables *LDL*, *CETP*, *IA* y *R*, respectivamente. La expresión (7) define una implicación del tipo Mamdani (Driankov *et al.*,

² Puede verse que el operador "y" es equivalente en este contexto al operador *mínimo*.

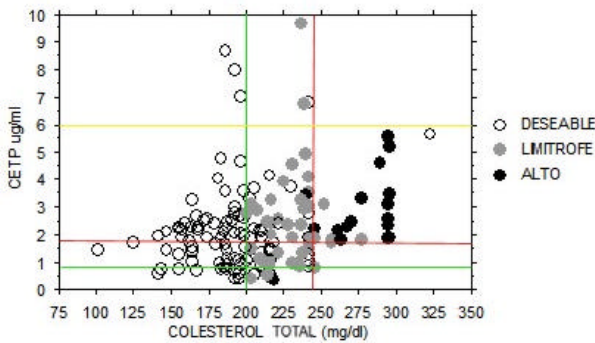


Figura 3. Mediciones de LDL clasificadas con respecto a *CT* y *CETP*. Datos tomados de una población de 162 sujetos al azar

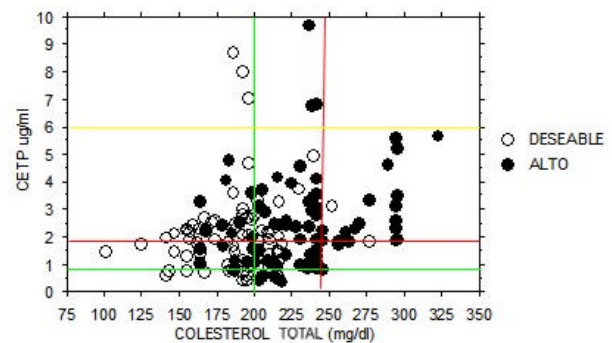


Figura 4. Mediciones de IA clasificadas con respecto a *CT* y *CETP*. Datos tomados de una población de 162 sujetos al azar

1993). Este tipo de implicación se eligió para diseñar el sistema de inferencias, pues el consecuente está dominado por el mínimo valor de todas las variables antecedentes. Esto equivale a decir que los valores pequeños de un factor, reducen la posibilidad del consecuente, independientemente del valor de los otros factores en juego. Esta forma de implicación coincide con la forma en que la comunidad médica maneja la combinación de un conjunto de signos en la formulación de un diagnóstico.

Los diferentes valores borrosos μ_m son el resultado de aplicar cada una de las diferentes reglas de inferencia con los posibles valores que *CT*, *LDL*, *CETP* e *IA* pueden llegar a tener y que inciden en el valor consecuente de *R*.

Los resultados individuales de las implicaciones μ_m deberán agregarse para tener un solo conjunto de valores obtenidos a través de la combinación de los resultados de las reglas activadas,

$$\mu_s(R) = \max\{\mu_1(R), \mu_2(R), \dots, \mu_n(R)\} \quad (8)$$

donde *n* es el número total de reglas y μ_s es el estado borroso del sistema de inferencias.

Para entender de manera gráfica el funcionamiento del sistema de inferencias y de la agregación de valores de pertenencia, considérese la

figura 5. Esta figura contiene 2 de las 24 reglas de inferencia propuestas³.

Suponga ahora que después de practicar el análisis de plasma a un paciente, se obtienen los valores de *CT_o*, *LDL_o*, *IA_o* y *CETP_o*. El sistema de inferencias está representado gráficamente en los dos primeros renglones de la figura 5. En la parte izquierda de esta figura se toman estos valores específicos, *CT_o*, *LDL_o*, *IA_o* y *CETP_o*, y se les proyecta sobre las funciones de pertenencia de los conjuntos borrosos antecedentes de las reglas 2 y 5. Para la regla 2, en el renglón superior, estas funciones corresponden con *CT_{deseable}*, *LDL_{deseable}*, *IA_{deseable}* y *CETP_{altoriesgo}*, para la regla 5, en el renglón inferior, son *CT_{deseable}*, *LDL_{deseable}*, *IA_{altoriesgo}* y *CETP_{altoriesgo}*. Las áreas sombreadas en las funciones de pertenencia del lado izquierdo se obtienen al sombreado el área que está por debajo de la ordenada o valor de pertenencia, que corresponde con los valores específicos *CT_o*, *LDL_o*, *IA_o* y *CETP_o*. Las líneas horizontales que atraviesan ambas reglas, toman la mínima ordenada de los conjuntos en la parte izquierda y la extienden hasta la parte derecha de la figura 5, que contiene las funciones de pertenencia de los conjuntos borrosos en los consecuentes de las reglas 2 y 5.

³ Se seleccionaron arbitrariamente las reglas 2 y 5 de las 24 reglas descritas en el apéndice de este artículo.

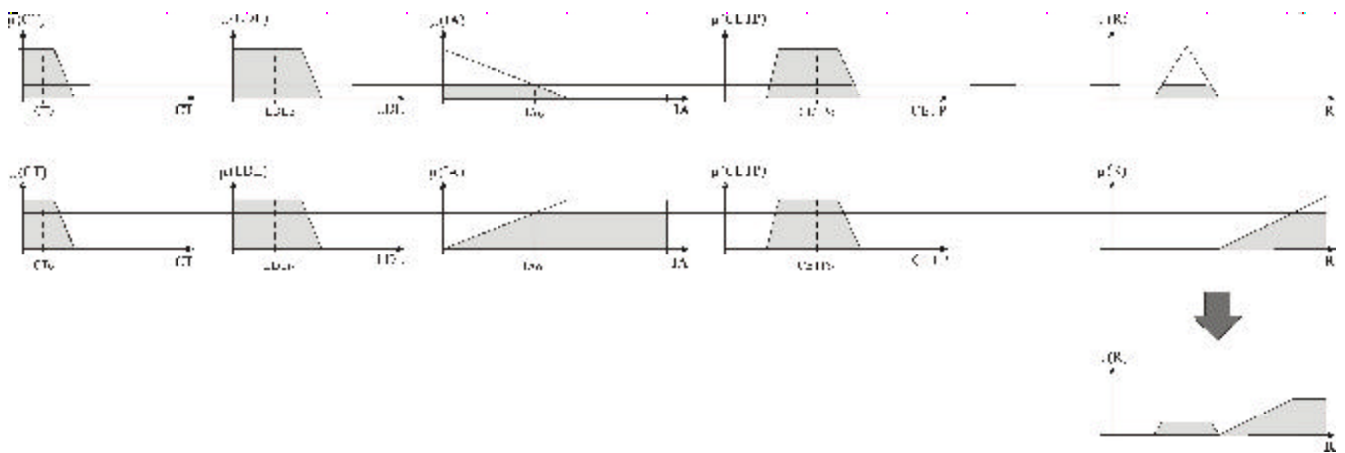


Figura 5. Disparo de reglas y agregación para una implicación del tipo Mamdani

Nótese que estas líneas horizontales definen el área sombreada en los consecuentes. Finalmente, la operación de agregación se muestra en el renglón inferior de la figura 5. Es claro que dicho renglón contiene la unión de las áreas sombreadas en el lado derecho de los dos primeros renglones.

Para traducir este valor borroso μ_s a un número real que represente el resultado buscado, se propone un método concretador⁴ al que se conoce como *el medio de la función de pertenencia máxima* (Ross, 1995).

$$\bar{R} = \frac{\arg_{\min}\{\max_{\forall p}\mu_p(R)\} + \arg_{\max}\{\max_{\forall p}\mu_p(R)\}}{2} \quad (9)$$

donde \arg_{\min} y \arg_{\max} denotan el argumento mínimo y máximo de la función $\max_{\forall p}\mu_p(R)$, respectivamente. Lo que se persigue es obtener la función de membresía en el consecuente, que después de haber realizado las inferencias, proporcione el máximo valor de riesgo y promedie los valores mínimo y máximo del dominio de dicha función de membresía. Este método de concreción se seleccionó, pues permite detectar para una combinación dada de valores de *CT*, *LDL*, *CETP* e *IA*, aquel consecuente que produce el máximo valor de riesgo.

La figura 6, que recupera el tercer renglón de la figura 5, ilustra el proceso de concreción. Puede observarse que el área sombreada a la derecha tiene una ordenada mayor que el área sombreada a la izquierda.

El valor señalado como R_{\min} corresponde con $\arg_{\min}\{\max_{\forall p}\mu_p(R)\}$ y el valor señalado con R_{\max} con $\arg_{\max}\{\max_{\forall p}\mu_p(R)\}$. Finalmente, el valor R_o es el promedio aritmético de R_{\min} y R_{\max} .

Como se ha mencionado, este trabajo parte de las investigaciones reportadas en Alonso (2003), Alonso *et al* (2003) y Alonso y Mas-Oliva (2002). El sistema de inferencia borrosa se diseñó para reproducir, de la mejor manera posible, los

hallazgos de estas investigaciones. Las funciones de pertenencia de cada parámetro y los operadores empleados fueron ajustados para reproducir los resultados clínicos. Existen claramente, otras opciones para definir la forma de estas funciones y operadores. La forma elegida hizo énfasis en su simplicidad y buscó aumentar la posibilidad de entendimiento del sistema de diagnóstico por parte de médicos no familiarizados con la lógica borrosa. Por esta misma razón, tampoco se recurrió a algoritmos para el reconocimiento de patrones en datos, como los sugeridos en Bezdek (1981) y Bezdek *et al.* (1999), pues en este caso, los rangos para la interpretación de los valores de los distintos factores no pueden moverse libremente como resultado de algún proceso de agrupación, lo que resta interés en las técnicas citadas.

Resultados

El sistema de inferencias borrosas diseñado, puede interpretarse como un mapeo $\mathfrak{R}^4 \rightarrow \mathfrak{R}$, o más precisamente como una función $R = f(CT, LDL, CETP, IA)$. Como ejemplo de su aplicación directa se muestra la tabla 3, que contiene los resultados obtenidos al aplicar el sistema a 10 individuos. Cada renglón contiene los valores de *CT*, *LDL*, *CETP* e *IA*, resultado del análisis de plasma sanguíneo para cada individuo y el nivel de riesgo de aterogénesis calculado por el sistema de inferencias. Puede notarse claramente que dos personas se clasificaron como *SinRiesgo* tres como *BajoRiesgo* y las cinco restantes como *AltoRiesgo*, según los rangos indicados en la figura 2.

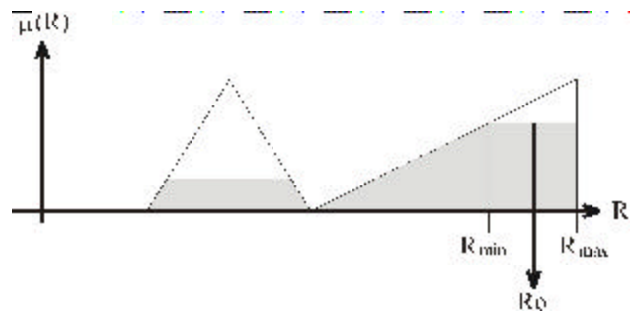


Figura 6. Método concretador. Promedio de los máximos

⁴ El término concretar es usado aquí como la operación que convierte a un número borroso en uno real, debido a que concreto es lo opuesto a difuso o borroso.

Tabla 3. Valores de CT, LDL, CETP e IA de distintos sujetos sometidos a análisis sanguíneo y riesgo de aterogénesis asociado a cada uno, a través del sistema de inferencias borrosas

CT [75,600]	LDL [50,250]	CETP [0,10]	IA [0,16]	RIESGO [1,100]
200	101.2	1.1	4.4	45.7
245	140.6	1.9	5.2	88.8
187	107.2	0.5	4.5	34.9
210	129.2	0.95	4.2	43.8
241	119.4	1.1	4.5	70.3
210	129.2	1.9	4.2	68.9
168	91	2.3	4.7	71.6
173	83.8	1.93	2.4	59.9
322	116	5.7	7.9	94.4
140	55.6	0.2	2.0	22.2

El análisis detallado de la tabla 3 revela con claridad la influencia del factor *CETP* en el diagnóstico formulado. Los resultados marcados para el tercero, y sexto a octavo renglones, no se podrían haber obtenido de utilizarse únicamente la tabla 1. Por otro lado, en el resto de los casos, el uso del *CETP* permitió corroborar el riesgo del paciente.

Como se mencionó en la introducción, el segundo objetivo principal de esta investigación fue proporcionar a los investigadores y médicos una herramienta de fácil uso para poder calibrar cambios o realizar adiciones en un futuro a este sistema de diagnóstico o alguno similar. En este

sentido, el sistema de inferencias diseñado permite obtener gráficas de mapeos $\mathfrak{R}^2 \rightarrow \mathfrak{R}$ que facilitan la interpretación de los resultados. La figura 7 muestra la superficie que se obtiene al graficar $R = f_1(CT, CETP)$, esto es, la superficie que se obtiene al dejar a *CT* y *CETP* como variables independientes y mantener fijas a $LDL = 150 \frac{\mu g}{ml}$ e $IA = 4.2$ ⁵; la figura 8 muestra esa misma superficie desde un plano superior.

⁵ Estos valores corresponden con regiones de transición para *LDL* e *IA*.

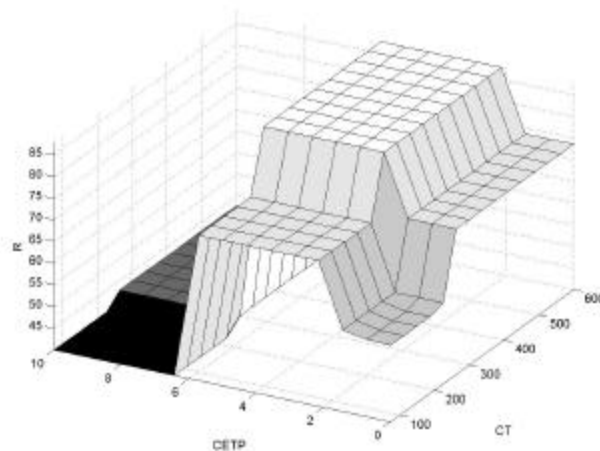


Figura 7: Superficie obtenida de graficar $(CT, CETP) \rightarrow R$

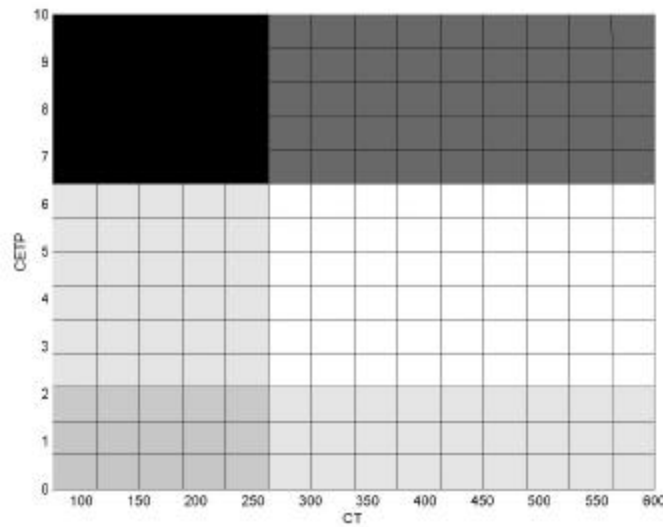


Figura 8. Vista de planta de la superficie $(CT, CETP) \rightarrow R$ donde pueden apreciarse distintas zonas de riesgo (negro \Rightarrow Sin Riesgo, blanco \Rightarrow Alto Riesgo).

La figura 8 se puede comparar con los datos clínicos presentados en las figuras 3 y 4, obtenidas de mediciones clínicas. Nótese que en el caso de la figura 3 estas zonas de riesgo están definidas con respecto a los valores e intervalos en los que se ha categorizado a la variable LDL en forma clínica. Por otra parte, la figura 4 también muestra zonas de riesgo establecidas al graficar valores clínicos de CT y $CETP$, con respecto al Índice Aterogénico IA . Las zonas de riesgo de ambas figuras son análogas entre sí y lo son con las de la figura 8.

Aunque en la figura 8 es posible apreciar las categorías definidas para CT ($CT_{deseable}$ y $CT_{alto riesgo}$) y las establecidas para $CETP$ ($CETP_{deseable}$, $CETP_{alto riesgo}$ y $CETP_{sobreexpresión}$), es en la figura 7 donde son visibles las transiciones entre cada una de ellas, especialmente aquellas regiones de transición que en la tabla 1 o en la figura 3 corresponden con el conjunto de valores denominado *límitrofe*.

Así, es posible hacer que el sistema de inferencias diseñado auxilie en la visualización de otras combinaciones de variables de entrada tomadas como independientes, si se fijan las demás. En esta forma, el médico puede obtener representaciones de todas las variables involucradas en una región definida de trabajo y hacerse una idea del comportamiento al variar cantidades de

interés. También es posible modificar con facilidad las funciones de pertenencia y observar las consecuencias de ello en el diagnóstico de riesgo.

Las figuras 3 y 4 son útiles en el diagnóstico clínico habitual. Sin embargo, cuando se quiere hacer el diagnóstico completo, el profesional debe establecer relaciones entre las cuatro variables en forma simultánea, tal y como lo hacen cada una de las reglas del sistema mostrado en este trabajo. Es claro que no será simple establecer estas relaciones a medida que el número de variables, o sus particiones, aumenten. Así, la utilidad de un sistema de inferencias como el propuesto resulta clara.

Conclusiones

Se presentó el diseño de un sistema de inferencias borrosas que auxilia en el diagnóstico del riesgo de aterogénesis, un padecimiento con alta incidencia en nuestro país y de graves consecuencias para las personas que lo padecen.

El sistema se basa en los resultados de un análisis simple de plasma sanguíneo que detecta los niveles de Colesterol Total (CT), Lipoproteínas de Baja Densidad (LDL) y calcula el Índice Aterogénico (IA). Se añade, por primera vez, la medición de un nuevo factor, la cantidad de Proteína Transferidora de Ésteres de Colesterol ($CETP$).

A partir de resultados de investigaciones clínicas, se propone un sistema de inferencias borrosas que permite determinar de manera cuantitativa el riesgo de aterogénesis en un paciente, dados los niveles de los cuatro factores detectados en el análisis de plasma sanguíneo. La asistencia de este sistema en el diagnóstico clínico ayuda al profesional a tomar decisiones a partir de datos obtenidos en forma habitual.

Por otro lado, el sistema aquí diseñado no es estático y está sujeto a mejoras. Se ha proporcionado a los investigadores que lo utilizan un mecanismo para ajustar con facilidad las reglas, funciones de pertenencia y la forma para realizar las agregaciones y concreciones. Con ello, será posible incorporar resultados de nuevas investigaciones o cambios surgidos de consultas a la comunidad médica sobre los valores y categorías establecidas para cada variable aquí utilizada. Un ejemplo claro es el caso de la variable *CETP*, hasta ahora no tomada en cuenta por dicha comunidad médica y que se propone como elemento importante para elaborar diagnósticos más precisos.

Referencias

- Alonso A.L. (2003). Identificación de una nueva isoforma de la proteína transferidora de ésteres de colesterol y evaluación de los niveles de CETP plasmática en población mexicana. Tesis de doctorado en biotecnología, Facultad de Química, UNAM.
- Alonso A.L., Dehesa A.Z. and Mas-Oliva J. (2003). Characterization of a Naturally Occurring New Version of the Cholesterol Ester Transfer Protein (CETP) From Small Intestine. *Molecular and Cellular Biochemistry*, K.A. Publishers, Vol. 245, pp. 173-182.
- Alonso A.L. and Mas-Oliva J. (2002). Immunoenzymatic Quantification Method. En: *WO 02/92995 A1*.
- Bezdek J.C. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms*. New York, Plenum.
- Bezdek J.C., Keller J.M., Krishnapuram R. and Pal N.R. (1999). *Fuzzy Models and Algorithms for Pattern Recognition and Image Processing*. Norwell, MA, Kluwer.
- Choi Y. and Krishnapuram R. (1995, 9-10 June). A Fuzzy-Rule-Based Image Enhancement Method for Medical Applications. Proceedings of the 8th IEEE Symposium on Computer-Based Medical Systems, pp. 75-80.
- Driankov D., Hellendoorn H. and Reinfrank M. (1993). *An Introduction to Fuzzy Control*. USA, Springer-Verlag.
- Gorzalczany M.B. and McLeish M. (1992, June 14-17). Combination of Neural Networks and Fuzzy Sets as a Basis for Medical Expert Systems. Proceedings of the Fifth Annual IEEE Symposium on Computer-Based Medical Systems, Durham, NC, USA, pp. 412-420.
- Hudson D.L. and Cohen M.E. (1994, November-December). Fuzzy Logic in Medical Expert Systems. *IEEE Engineering in Medicine and Biology* 13(5), 693-698.
- Morato H.L. y Posadas R.C. (1996). *Atención integral del paciente diabético en diagnóstico y tratamiento de las dislipidemias*. G.I. Lerman, editorial Interamericana-McGraw Hill, México.
- Nguyen H.P., Nguyen B.T., Ding L. and Hirota K. (2001, July). Case Based Reasoning Using Fuzzy Set Theory and the Importance of Features in Medicine. *Proceedings of the IFSA World Congress and 20th NAFIPS International Conference*, Vol. 2, pp. 872-876. 9th joint.
- Ross T.J. (1995). *Fuzzy Logic with Engineering Applications*. McGraw-Hill.
- Sidaoui H. and Pacheco M.T.T. (1999, 31st August-1st September). Fuzzy Algorithms in Medical Decision Making Processes. *Proceedings of the Third International Conference on Knowledge-Based Intelligent Information Engineering Systems*, Adelaide, Australia, pp. 300-304.
- Tseng H.C. and Teo D.W. (1994, June, 26-29). Medical Expert System with Elastic Fuzzy Logic. En: Proceedings of the Third IEEE Conference on Fuzzy Systems, Vol. 3, Orlando, FL, USA, pp. 2067-2071.
- Udupa J.K., Wei L., Samarasekera S., Miki Y., Van Buchem M.A. and Grossman R.I. (1997, October). Multiple Sclerosis Lesion Quantification Using Fuzzy-Connectedness Principles. *IEEE Transactions on Medical Imaging*, 16(5), 598-609.

Reglas de inferencia

1. si *CT* es *CTdeseable* y *LDL* es *LDLdeseable* y *CETP* es *CETPdeseable* e *IA* es *IAdeseable* entonces *Res SinRiesgo*
2. si *CT* es *CTdeseable* y *LDL* es *LDLdeseable* y *CETP* es *CETPaltorriesgo* e *IA* es *IAdeseable* entonces *Res BajoRiesgo*
3. si *CT* es *CTdeseable* y *LDL* es *LDLdeseable* y *CETP* es *CETPsobreexpresión* e *IA* es *IAdeseable* entonces *Res SinRiesgo*
4. si *CT* es *CTdeseable* y *LDL* es *LDLdeseable* y *CETP* es *CETPdeseable* e *IA* es *IAaltorriesgo* entonces *Res BajoRiesgo*
5. si *CT* es *CTdeseable* y *LDL* es *LDLdeseable* y *CETP* es *CETPaltorriesgo* e *IA* es *IAaltorriesgo* entonces *Res AltoRiesgo*
6. si *CT* es *CTdeseable* y *LDL* es *LDLdeseable* y *CETP* es *CETPsobreexpresión* e *IA* es *IAaltorriesgo* entonces *Res BajoRiesgo*
7. si *CT* es *CTdeseable* y *LDL* es *LDLaltorriesgo* y *CETP* es *CETPdeseable* e *IA* es *IAdeseable* entonces *Res BajoRiesgo*
8. si *CT* es *CTdeseable* y *LDL* es *LDLaltorriesgo* y *CETP* es *CETPaltorriesgo* e *IA* es *IAdeseable* entonces *Res AltoRiesgo*
9. si *CT* es *CTdeseable* y *LDL* es *LDLaltorriesgo* y *CETP* es *CETPsobreexpresión* e *IA* es *IAdeseable* entonces *Res NoObservado*
10. si *CT* es *CTdeseable* y *LDL* es *LDLaltorriesgo* y *CETP* es *CETPdeseable* e *IA* es *IAaltorriesgo* entonces *Res AltoRiesgo*
11. si *CT* es *CTdeseable* y *LDL* es *LDLaltorriesgo* y *CETP* es *CETPaltorriesgo* e *IA* es *IAaltorriesgo* entonces *Res AltoRiesgo*
12. si *CT* es *CTdeseable* y *LDL* es *LDLaltorriesgo* y *CETP* es *CETPsobreexpresión* e *IA* es *IAaltorriesgo* entonces *Res AltoRiesgo*
13. si *CT* es *CTaltorriesgo* y *LDL* es *LDLdeseable* y *CETP* es *CETPdeseable* e *IA* es *IAdeseable* entonces *Res BajoRiesgo*
14. si *CT* es *CTaltorriesgo* y *LDL* es *LDLdeseable* y *CETP* es *CETPaltorriesgo* e *IA* es *IAdeseable* entonces *Res AltoRiesgo*
15. si *CT* es *CTaltorriesgo* y *LDL* es *LDLdeseable* y *CETP* es *CETPsobreexpresión* e *IA* es *IAdeseable* entonces *Res NoObservado*
16. si *CT* es *CTaltorriesgo* y *LDL* es *LDLdeseable* y *CETP* es *CETPdeseable* e *IA* es *IAaltorriesgo* entonces *Res AltoRiesgo*
17. si *CT* es *CTaltorriesgo* y *LDL* es *LDLdeseable* y *CETP* es *CETPaltorriesgo* e *IA* es *IAaltorriesgo* entonces *Res AltoRiesgo*
18. si *CT* es *CTaltorriesgo* y *LDL* es *LDLdeseable* y *CETP* es *CETPsobreexpresión* e *IA* es *IAaltorriesgo* entonces *Res AltoRiesgo*
19. si *CT* es *CTaltorriesgo* y *LDL* es *LDLaltorriesgo* y *CETP* es *CETPdeseable* e *IA* es *IAdeseable* entonces *Res AltoRiesgo*
20. si *CT* es *CTaltorriesgo* y *LDL* es *LDLaltorriesgo* y *CETP* es *CETPaltorriesgo* e *IA* es *IAdeseable* entonces *Res AltoRiesgo*
21. si *CT* es *CTaltorriesgo* y *LDL* es *LDLaltorriesgo* y *CETP* es *CETPsobreexpresión* e *IA* es *IAdeseable* entonces *Res NoObservado*
22. si *CT* es *CTaltorriesgo* y *LDL* es *LDLaltorriesgo* y *CETP* es *CETPdeseable* e *IA* es *IAaltorriesgo* entonces *Res AltoRiesgo*
23. si *CT* es *CTaltorriesgo* y *LDL* es *LDLaltorriesgo* y *CETP* es *CETPaltorriesgo* e *IA* es *IAaltorriesgo* entonces *Res AltoRiesgo*
24. si *CT* es *CTaltorriesgo* y *LDL* es *LDLaltorriesgo* y *CETP* es *CETPsobreexpresión* e *IA* es *IAaltorriesgo* entonces *Res AltoRiesgo*

Semblanza de los autores

Oscar A. Rosas-Jaimes. Actualmente es estudiante de doctorado en ingeniería eléctrica en la UNAM. Es ingeniero mecánico y maestro en ingeniería eléctrica en el campo del control automático, ambos por la UNAM. Ha sido profesor de los laboratorios de mecánica en la Facultad de Ingeniería de la UNAM y del curso propedéutico de geometría y trigonometría en la misma institución. Sus áreas de interés son el control y estimación de tráfico vehicular y el uso de sistemas de lógica borrosa en el diagnóstico médico.

Ana L. Alonso-García. Obtuvo su licenciatura en Biología y su doctorado en Biotecnología en la UNAM. Realiza una estancia posdoctoral en la Unidad de Investigación Médica Bioquímica del Hospital de Especialidades, Centro Médico Nacional Siglo XXI, a cargo del Dr. Miguel Cruz López. Es responsable del Banco Nacional de DNA dentro de un proyecto nacional para la identificación de genes asociados a Diabetes Tipo II. Fue profesora de la maestría en ciencias en biotecnología genómica del IPN. Recibió el Premio "Dr. Jesús Kumate Rodríguez" y el premio CANIFARMA en el área de investigación tecnológica. Es titular de varias patentes.

Jaime Mas-Oliva. Es investigador titular del Instituto de Fisiología y coordinador del Programa Universitario de Investigación en Salud, ambos de la UNAM. Es médico cirujano por la Facultad de Medicina, UNAM y doctor en bioquímica por el National Heart and Lung Institute, Imperial College for Science, Technology and Medicine, London University, UK. Es investigador nacional nivel III. Ha recibido diversas distinciones entre las que destacan el Premio de la Academia Mexicana de Ciencias, Distinción Universidad Nacional Autónoma de México, Premio Manuel Noriega Morales de la OEA, Premio Canifarma, Beca John Simon Guggenheim, Medal of Merit, Institute for Cardiovascular Sciences. Ha dirigido 27 tesis, publicado 98 artículos de circulación internacional con 856 citas y es autor de 4 libros.

Luis Álvarez-Icaza. Es investigador titular de la Coordinación de Automatización del Instituto de Ingeniería de la UNAM. Imparte cátedra en la División de Ingeniería Eléctrica de la Facultad de Ingeniería, así como en los posgrados de ingeniería y ciencias de la computación, todos de la UNAM. Obtuvo su licenciatura en ingeniería electromecánica y su maestría en ingeniería en la Facultad de Ingeniería de la UNAM. Cursó sus estudios de doctorado en ingeniería mecánica en la Universidad de California en Berkeley. Sus líneas de investigación se refieren principalmente al control de sistemas no-lineales con aplicaciones al control de tráfico vehicular, control avanzado de vehículos y control de estructuras civiles.