

Árbol para predecir el desarrollo de la cardiopatía hipertensiva

Prediction tree for hypertensive heart disease

Dr. Alexis Álvarez Aliaga,^I Dr. Julio César González Aguilera,^I Dra. Liliana del Rosario Maceo Gómez,^{II} Dr. Adonis Frómata Guerra,^I Dr. Salvador Bázaga Morell,^I Dra. Ana Elis Cervantes Infante^I

^I Hospital General Universitario "Carlos Manuel de Céspedes". Bayamo, Cuba.

^{II} Policlínico "Jimmy Hirzel". Bayamo, Cuba.

RESUMEN

Introducción: la cardiopatía hipertensiva representa altas morbilidad y mortalidad. La complejidad de su prevención radica en el número de factores de riesgo implicados en su desarrollo.

Objetivo: diseñar y validar un árbol de predicción para el desarrollo de la cardiopatía hipertensiva, a partir de factores hemodinámicos y no hemodinámicos.

Métodos: se realizó el diseño y validación de un árbol de predicción para el desarrollo de la cardiopatía hipertensiva, mediante el procedimiento de *descubrimiento de conocimientos en bases de datos y minería de datos*, en pacientes hipertensos atendidos en la consulta especializada de hipertensión arterial de la Policlínica de Especialidades del Hospital General Universitario "Carlos Manuel de Céspedes" del municipio Bayamo, provincia de Granma, Cuba, desde el 1^o de enero de 2004 hasta el 31 de diciembre de 2009.

Resultados: el árbol predijo el riesgo de desarrollar la cardiopatía hipertensiva a 82,598 % de los pacientes; con un área bajo la curva ROC de 0,861 y una tasa de verdaderos positivos de 0,733 y de 0,921 para las clases 1 y 2, respectivamente. El factor más importante lo constituyó la proteína C reactiva, seguida en orden de importancia por la glucemia, el ácido úrico, el colesterol y la microalbuminuria.

Conclusiones: el árbol de toma de decisiones y el conjunto de reglas derivado de este, permitieron predecir el riesgo de desarrollar la cardiopatía hipertensiva en individuos hipertensos, por lo que su introducción en la práctica clínica, mejorará la evaluación del paciente con hipertensión arterial.

Palabras clave: cardiopatía hipertensiva, hipertensión arterial, factores de riesgo, árbol de predicción.

ABSTRACT

Introduction: hypertensive heart disease is associated with high morbidity and mortality. The complexity of its prevention lies in the number of risk factors involved in its development.

Objective: design and validate a prediction tree for the development of hypertensive heart disease on the basis of hemodynamic and non-hemodynamic factors.

Methods: a prediction tree for the development of hypertensive heart disease was designed and validated following the procedure of knowledge discovery in databases and data mining. The patients surveyed were being cared for at the hypertension specialized consultation of the secondary care polyclinic in Carlos Manuel de Céspedes General University Hospital, municipality of Bayamo, Granma province, Cuba, from 1 January 2004 to 31 December 2009.

Results: the tree predicted risk of developing hypertensive heart disease in 82.598% of the patients, with an area of 0.861 under the ROC curve and a true positives rate of 0.733 and 0.921 for classes 1 and 2, respectively. The most important factor was C-reactive protein, followed by glycemia, uric acid, cholesterol and microalbuminuria, in that order.

Conclusions: the decision making tree and the set of rules deriving from it allowed prediction of the risk to develop hypertensive heart disease in high blood pressure patients. Their incorporation into clinical practice will improve the evaluation of these patients.

Key words: hypertensive heart disease, arterial hypertension, risk factors, prediction tree.

INTRODUCCIÓN

Desde la segunda mitad del pasado siglo XX, las enfermedades cardiovasculares (ECV) constituyen una de las principales causas de muerte en los países desarrollados y en vías de desarrollo.

Entre las ECV, la hipertensión arterial (HTA) posee una alta prevalencia en la población en general y, en particular, después de los 60 años de edad en individuos de ambos géneros, constituye uno de los factores de riesgo de mayor importancia para la aparición de enfermedades del corazón, por lo que su control se considera entre las medidas fundamentales para evitarlas.¹

La HTA no puede ser considerada una entidad aislada, por su evolución natural, incrementa considerablemente el riesgo de lesión a diferentes órganos, así como la invalidez y las cifras muerte por causa cardiovascular.^{1,2}

La afección a órganos diana por HTA que presenta mayores morbilidad y mortalidad es la cardiopatía hipertensiva. Se define como un complejo y variable conjunto de efectos que provoca en el corazón la elevación crónica de la presión arterial y se caracteriza por la presencia de signos anatómicos³ o bioquímicos^{4,5} de hipertrofia ventricular izquierda (HVI) o disfunción ventricular diastólica o sistólica, de isquemia miocárdica y de alteraciones del ritmo cardíaco.⁶

Sin embargo, a pesar de la importancia que tiene en el desarrollo de la hipertrofia de los cardiomiocitos, la sobrecarga hemodinámica que ocasiona la HTA, solo justifica el 50 % de la lesión orgánica; lo cual indica que un número importante de factores de riesgo y procesos patológicos mediados por hormonas, factores de crecimiento, citocinas y otras moléculas inflamatorias, al accionar sobre el cardiomiocito y el resto de los componentes celulares y no celulares del miocardio, pueden explicar el desarrollo de la cardiopatía hipertensiva en el paciente hipertenso.⁶

Estos factores pueden emplearse en la construcción de herramientas para predecir el riesgo cardiovascular, que son escalas con las cuales se trata de unir toda la información de un enfermo donde se le otorga a cada factor un valor, según su importancia, para lograr un número que permita al médico de asistencia tener una idea general del riesgo del paciente y actuar en consecuencia.

Basado en lo expresado anteriormente, la presente investigación tiene como objetivo diseñar y validar un árbol de predicción para el desarrollo de la cardiopatía hipertensiva, a partir de factores hemodinámicos y no hemodinámicos.

MÉTODOS

Diseño metodológico

Se realizó el diseño y validación de un árbol de predicción del desarrollo de la cardiopatía hipertensiva mediante el procedimiento de *descubrimiento de conocimientos en bases de datos*, conocido internacionalmente como proceso KDD (del inglés *Knowledge Discovery in Database*) y *minería de datos (Data Mining- DM)*, en pacientes hipertensos atendidos en la consulta especializada de HTA de la Policlínica de Especialidades del Hospital General Universitario "Carlos Manuel de Céspedes" del municipio Bayamo, provincia Granma, Cuba, desde el 1^o de enero de 2004 hasta el 31 de diciembre de 2009.

Criterios de inclusión

- Pacientes hipertensos de 20 años o más de edad.
- Pacientes con antecedente patológico personal de HTA esencial de 10 años o más, tiempo en que es más frecuente el daño orgánico.^{1,2,7,8-11}

Criterios de exclusión

- Adolescentes con HTA esencial (por la poca frecuencia de daño a órganos diana).
- Paciente con HTA esencial de menos de 10 años de evolución.

- Pacientes con cardiopatía isquémica, a pesar de su elevada frecuencia en el hipertenso, donde el rol de la HTA, aunque evidente, no sería el único factor influyente en su aparición (en la presente investigación se evaluaron los efectos directos de la HTA, por lo que la inclusión de esta forma clínica puede inducir a sesgos de selección y clasificación).
- Enfermos con trastornos de la conducción interventricular y auriculoventricular.
- Pacientes que no padecieran otros estados mórbidos que pudieran provocar la cardiopatía.

Casos (incidentes)

Se definieron aquellos pacientes hipertensos, según *The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation and Treatment of High Blood Pressure*,¹ con 1 o más de los siguientes elementos diagnósticos de cardiopatía hipertensiva, en los cuales no se encontró otra causa que justificara la cardiopatía:

- Patrón ecocardiográfico de HVI: se aplicó la fórmula de Devereux¹² (masa del VI en gramos) = $0,8 (1,04 [\text{diámetro diastólico ventricular izquierdo} + \text{espesor pared posterior} + \text{espesor septum interventricular}]^3 - [\text{diámetro diastólico ventricular izquierdo}]^3) + 0,6$. Se consideró con hipertrofia al valor $\geq 125 \text{ g/m}^2$ en varones y de $\geq 110 \text{ g/m}^2$ en las mujeres.
- Disfunción diastólica (alteración en la distensibilidad diastólica, el llenado o relajación del ventrículo izquierdo) o elementos de disfunción sistólica (fracción de eyección $< 45 \%$), por ecocardiograma.^{6,13,14}
- Manifestaciones clínicas de insuficiencia cardíaca de reciente comienzo (primera presentación), tales como: historia de disnea de esfuerzo, disnea paroxística nocturna, estertores crepitantes bibasales, ritmo de galope por tercer tono y antecedentes de edema agudo del pulmón hipertensivo,^{6,15,16} sumado a uno o más de los hallazgos ecocardiográficos descritos antes.
- Arritmia: se consideró la fibrilación atrial persistente de larga evolución y la permanente no causada por las valvulopatías, miocardiopatías, miocarditis, fármacos y drogas. La confirmación se obtuvo con un electrocardiograma estándar de 12 derivaciones, de acuerdo con los criterios propuestos por el Colegio Americano de Cardiología y la Asociación Americana del Corazón.¹⁷ El trazado electrocardiográfico fue realizado por la enfermera de la consulta con electrocardiógrafo digital portátil (CARDIOCID-BB, modelo A5102).

El ecocardiograma fue realizado por 2 cardiólogos con más de 15 años de experiencia y especializados en ecocardiografía. Se utilizó un equipo ASAOTE Caris PLUS y se aplicaron las normativas de la Asociación Americana de Ecocardiografía.

Testigos

Pacientes hipertensos que no cumplieron los criterios clínicos, eléctricos y ecocardiográficos de cardiopatía hipertensiva antes definidos.

Características del universo y de la muestra

Fueron atendidos 4 857 pacientes en la consulta de HTA durante 5 años, procedentes de áreas urbanas (n= 3 067; 63,15 %) y rurales (n= 1 790; 36,85 %) de la provincia.

El tamaño de la muestra se determinó mediante el programa EPIDAT 3.0. Para su estimación se consideraron los parámetros siguientes: nivel de confianza al 95 %, poder del estudio al 90 %, relación de casos y testigos 1:2, proporción de casos expuestos 52,80 % y proporción de testigos expuestos 35,87 % y odds ratio (OR: oportunidad relativa) mínimo a detectar, 2. Se obtuvo un número de pares de 183.

Los pacientes se seleccionaron de forma aleatoria simple (sin reposición de los elementos: cada elemento extraído se descarta para la extracción, 1 caso por cada 2 testigos (1:2) y emparejados por género.

Se decidió investigar una muestra de 1 200 pacientes (400 casos y 800 testigos). Debido al interés por crear y validar el árbol para predecir el riesgo de cardiopatía hipertensiva, se dividió la muestra de forma aleatoria en 2 partes: una (66 % del total de muestra) para crear el árbol y la otra (33 % restante), para autovalidación, mediante la herramienta *Waikato Environment for Knowledge Analysis* (WEKA) v3.6.1 del software libre programado en JAVA.

Métodos empleados en la investigación

- Del nivel teórico: método histórico-lógico, de análisis-síntesis, hipotético-deductivo, sistémico-funcional-estructural, inducción-deducción, método holístico-dialéctico.
- Del nivel empírico: la observación, selección y consulta a expertos, la encuesta y la entrevista.
- Método estadístico-analítico, para determinar y evaluar los resultados.

Delimitación y operacionalización de las variables

Variable dependiente, fue la cardiopatía hipertensiva (la cual fue dicotomizada, en clase 1: presente y clase 2: ausente); variables independientes, aquellos factores cuya influencia en la aparición de la cardiopatía hipertensiva se estaba evaluando.

La edad se cuantificó en años cumplidos, variable que se convirtió en dicotómica, al considerar como expuestos a los pacientes mayores de 65 años.

El hábito de fumar fue agrupado en 2 categorías: fumadores (expuestos) si consumían diariamente o casi a diario cigarrillos, tabacos o pipa independiente del número fumado, y exfumadores de menos de un año, y como no fumadores (no expuesto), al resto de los pacientes.

Como alcoholismo se consideró a la ingestión mayor de 1 onza de alcohol puro diario, que equivale a 1 onz (20 mL) de etanol, 8 onz (240 mL) de vino, 24 onz (720 mL) de cerveza, 1½ onz (45 mL) de ron, en el caso de las mujeres y del bajo peso se reduce a 15 mL/d.

La obesidad fue establecida según el cálculo del índice de masa corporal ($IMC \geq 30$, peso en kg/talla en m^2) o una circunferencia de la cintura ≥ 102 cm para los hombres y 88 cm para las mujeres o ambas condiciones.

Se consideró como sedentaria a la persona que invierte diariamente (o casi a diario) menos de 25 (mujeres) o 30 (hombres) minutos en actividades de ocio (cualquier ejercicio de intensidad igual o superior a caminar a paso rápido) y también se tuvo en cuenta el tipo de profesión del paciente.

Como exceso de sal en la dieta se consideró la ingestión de sal > 5 g/d/por persona; esto equivale a más de una cucharadita de postre rasa por persona para cocinar, distribuida entre los platos confeccionados en el almuerzo y la comida.

Los marcadores biológicos que se seleccionaron como posibles factores de riesgo fueron el colesterol, ácido úrico, triglicéridos, HDL-colesterol, glucemia, PCR microalbuminuria y el cociente colesterol/HDL. Las muestras sanguíneas para cada uno de los exámenes de laboratorio se obtuvieron en ayunas (10 a 12 h) y fueron centrifugadas a temperatura ambiente a 2 000 rpm durante 10 min. La creatinina, ácido úrico, colesterol, HDL-colesterol, triglicéridos y glucemia, fueron medidos con el autoanalizador HITACHI 902® en las primeras 24 h después de su extracción. La determinación de todos los estudios se realizó por métodos enzimáticos.

La PCR no ultrasensible > 4,5 mg/L fue considerada como factor de riesgo hipotético y se determinó mediante el método de determinación cuantitativa turbidimétrico.¹⁸

Se consideró a los pacientes con microalbuminuria cuando sus cifras oscilaron desde 0,02 hasta 0,2 g/L en 24 h⁸ y se cuantificó mediante la técnica Microalb-Látex.¹⁹

Se consideraron puntos de corte para el análisis los valores siguientes: colesterol sérico por encima de 4,8 mmol/L, hipertrigliceridemia > 1,7 mmol/L, HDL-colesterol < 1,5 mmol/L el cociente colesterol/HDL > 3,5. El punto de corte para la glucemia se estableció cuando las cifras en ayunas alcanzaron valores superiores de 5,4 mmol/L, de 80 µmol/L para la creatinina y de 375 µmol/L para el ácido úrico.

El valor de las variables cuantitativas resultó del promedio de 3 resultados en las 3 primeras consultas. Se convirtieron en dicotómicas, teniendo en cuenta el punto que demostró mayor sensibilidad y especificidad, según el análisis de la curva característica operador receptor (*Receiver Operating Characteristics* -ROC); análisis realizado con el programa SPSS 17.0 para Windows.

Se definieron como pacientes controlados aquellos con cifras de presión arterial inferiores a 140 y 90 mmHg (sistólica y diastólica, respectivamente) en el 100 % de las tomas en consulta. El tiempo de evolución: los pacientes fueron agrupados en 2 categorías; pacientes con un tiempo de evolución entre 10 y 20 años, y pacientes con un tiempo de evolución > 20 años. El estadio de la HTA se clasificó según la propuesta del 7mo. informe de la NYHA.¹

Para la toma de la presión arterial se utilizaron esfigmomanómetros aneroides y de mercurio previamente calibrados por la Oficina Territorial de Normalización y Calidad (OTN), acreditada para tales fines.

Los datos se obtuvieron en las consultas durante las diferentes entrevistas realizadas por los autores, previo conocimiento y consentimiento del paciente.

Análisis estadísticos

Mediante el procedimiento KDD y DM, se diseñó un árbol de predicción para el desarrollo de la cardiopatía hipertensiva y un algoritmo diagnóstico autovalidado.

Para este propósito se utilizó la herramienta WEKA v3.6.1 del software libre programado en JAVA, que es un sistema que combina la técnica de KDD con DM.

Para poder procesar los datos, es necesario convertirlos a un formato de archivo especial, llamado ARFF.²⁰

Durante el desarrollo de este estudio se siguieron los pasos del proceso KDD que se describen a continuación, a través de los cuales se crearon las reglas y los árboles de clasificación, útiles para la predicción de la cardiopatía hipertensiva.

Creación de la base de datos de trabajo

Se creó una base de datos en el tabulador Microsoft Excel para lograr una mejor manipulación de los datos y quedó conformada con 19 atributos (variables operacionalizadas de forma dicotómica en presente y ausente), se incluyó la clase *cardiopatía hipertensiva*.

Después que se tuvieron los datos listos, como se mencionó anteriormente, se convirtió en formato de archivo ARFF, compuesto por 3 partes fundamentales (cabecera, declaración de atributos y sección de datos) y se utilizó el bloc de notas para editarlo.²¹

Reducción de los datos

La selección de atributos es una actividad muy importante dentro del proceso de KDD, por su utilidad para reducir la dimensionalidad del problema y con ello simplificar el clasificador, incrementar la velocidad de manipulación de los datos y mejorar el desempeño del clasificador con lo que se reduce la influencia de los ruidos. En la presente investigación se seleccionaron los atributos más relevantes. Para la selección del subconjunto de atributos se eligió como evaluador al algoritmo *CfsSubsetEval* y el método de búsqueda que se empleó fue la búsqueda exhaustiva. Este proceso tuvo como resultado que de los 19 atributos, que incluye a la clase, quedaran seleccionados solo 9 (aquellos que obtuvieron el 100 % de coincidencia en los 10 subconjunto de la muestra). Estos fueron: estadio de la HTA, tiempo de evolución de HTA, microalbuminuria, ácido úrico, creatinina, glucemia, proteína C reactiva y colesterol; la clase *cardiopatía hipertensiva* se mantuvo para los próximos pasos.

Selección de la función de la minería de datos y el algoritmo de minar los datos

Dentro de las funciones de la minería de datos se seleccionó la de predicción,²¹ pues se tienen bien definidos la clase y el objetivo primordial de la investigación.

En esta investigación, después de evaluar los algoritmos JRip, J48, Naive Bayes, IBk, Id3 y NNge, se decidió seleccionar al J48, por ser un clasificador que genera un árbol de decisión, lo que permite extraer reglas de predicción para todas las clases, además, se tuvieron en cuenta atributos que pueden estar presentes o no, indistintamente, para cada clase;²⁰ también fue el algoritmo con mejor sensibilidad y mayor área bajo la curva ROC.

Para este algoritmo en las opciones de prueba se utilizó la escisión de porcentajes (*Percentage Split*), se dividió el conjunto de datos en 2 (66 % para el diseño y 33 % para la validación).

Selección y consulta a expertos

Con el propósito de evaluar la validez de presentación y contenido del árbol de predicción, se realizó la consulta a expertos. De 18 posibles expertos examinados se seleccionaron 8, por el método cualimétrico de preferencia, sobre la base de la determinación de la competencia del experto y la voluntariedad de cooperar en esta investigación. Se utilizó la autovaloración de la investigación por el propio experto basándose en el coeficiente K, por obtener puntuaciones entre 0,7 y 1.

Cálculo del coeficiente K: $K = \frac{1}{2} (Kc + ka)$ para 18 posibles expertos.

Donde:

Kc: es el coeficiente de conocimiento que tenía el experto sobre los factores de riesgo y fisiopatología de la cardiopatía hipertensiva, calculado sobre la base de la valoración del propio experto en una escala de 0 (desconocimiento total) a 10 (conocimiento máximo) y multiplicado por 0,1.

Ka: es el coeficiente de argumentación de los criterios determinados como resultado de la suma de los puntos alcanzados a partir de una tabla patrón.

Se realizaron los cálculos en cada caso, se obtuvo el coeficiente de argumentación en cada experto y se escogieron los 8 de más alta competencia, todos oscilaron entre 0,7 y 1 de puntuación, competencia media o alta. Cada experto se consultó de manera individual y anónima a través de una guía elaborada para ello, basado en su conocimiento y experiencia en cardiopatía hipertensiva. Ellos se pronunciaron sobre el grado de cumplimiento de cada uno de los elementos que componían el árbol de predicción, según 3 posibilidades: nada, moderadamente o mucho.

Regulaciones éticas

En la presente investigación se cumplieron los preceptos éticos básicos de los procesos investigativos clínico-epidemiológicos. La dirección del hospital y el comité de ética del centro, dieron su aprobación.

Con el objetivo de aumentar la precisión y exactitud de los datos necesarios para la investigación se tuvieron en cuenta el control de los siguientes sesgos: de selección, clasificación, confusión, información y precisión.

RESULTADOS

En la tabla 1 se observa que la edad mayor de 65 años representó el 66,08 % de la casuística y hubo un discreto predominio de los enfermos del género masculino (50,75 %). La mayoría de los pacientes se clasificaron en el estadio 1 de la HTA (60,33 %) y más de la mitad de ellos estaban controlados (60 %).

Tabla 1. Caracterización de la muestra. Variables cualitativas

Variables	Categoría	No.	(%)
Edad (dicotómica)	≤ 65 años	407	33,92
	> 65 años	793	66,08
Sexo	Masculino	609	50,75
	Femenino	591	49,25
Estadio de la HTA	Estadio 1	724	60,33
	Estadio 2	476	39,67
Control de la HTA	Controlado	720	60,00
	No controlado	480	40,00
Microalbuminuria	Sí	421	35,08
	No	779	64,92
Hábito de fumar	Sí	511	42,58
	No	689	57,42
Alcoholismo	Sí	371	30,92
	No	829	69,08
Obesidad	Sí	445	37,08
	No	755	62,92
Sedentarismo	Sí	573	47,75
	No	627	52,25
Exceso de sal	Sí	467	38,92
	No	733	61,08

N= 1 200.

La mayor desviación estándar de los valores promedio lo tuvieron el ácido úrico (343,84; ±100,06) y la creatinina (79,17; ±25,33), como se observa en la tabla 2. Solo la edad no mostró diferencias significativas cuando se compararon los valores promedio entre los casos y los testigos ($p= 0,437$) de toda la muestra.

Tabla 2. Caracterización de la muestra. Variables cuantitativas

Variables	Mínimo	Máximo	Media	Desviación estándar	Mediana
Edad	21,00	89,00	57,8850	12,07794	59,00
Colesterol	2,00	9,80	4,7243	1,1632	4,60
HTA*	10,00	50,00	19,7683	8,8567	18,00
Proteína C reactiva	0,10	9,70	4,1358	1,8648	4,100
Ácido úrico	79,00	765,00	343,8494	100,0642	327,00
Colesterol/HDL	1,02	31,74	3,9729	3,1836	2,81
HDL	0,19	3,23	1,5306	0,5456	1,600
Creatinina	30,00	274,00	79,1777	25,3337	76,00
Glucemia	2,60	8,40	4,6953	1,0059	4,60
Triglicéridos	0,13	4,77	1,7182	0,7882	1,66

*HTA: tiempo de evolución de la hipertensión arterial.

N= 1 200.

En la tabla 3 se realizó la comparación entre los casos y los testigos de toda la muestra y se encontró que solo el alcoholismo no tuvo diferencias significativas ($p= 0,270$).

Tabla 3. Comparación entre los casos y testigos

Variables	Casos N= 400	Testigos N= 800	Total N= 1 200	*p
Proteína C reactiva	290	169	459	0,000
Colesterol	224	167	391	0,000
Creatinina	261	255	516	0,000
Glucemia	224	155	379	0,000
Ácido úrico	195	175	370	0,000
Evolución > 20 años	275	214	489	0,000
Microalbuminuria	241	185	426	0,000
Colesterol/HDL	245	176	421	0,000
Triglicéridos	234	313	547	0,000
No control de la HTA	259	221	480	0,000
Estadio 2 de la HTA	258	231	489	0,000
HDL	244	222	466	0,000
Edad > 65 años	216	170	386	0,000
Hábito de fumar	225	286	511	0,000
Obesidad	197	248	445	0,000
Sedentarismo	239	334	573	0,000
Exceso de sal	209	279	488	0,000
Alcoholismo	132	239	371	0,270

HTA: hipertensión arterial.

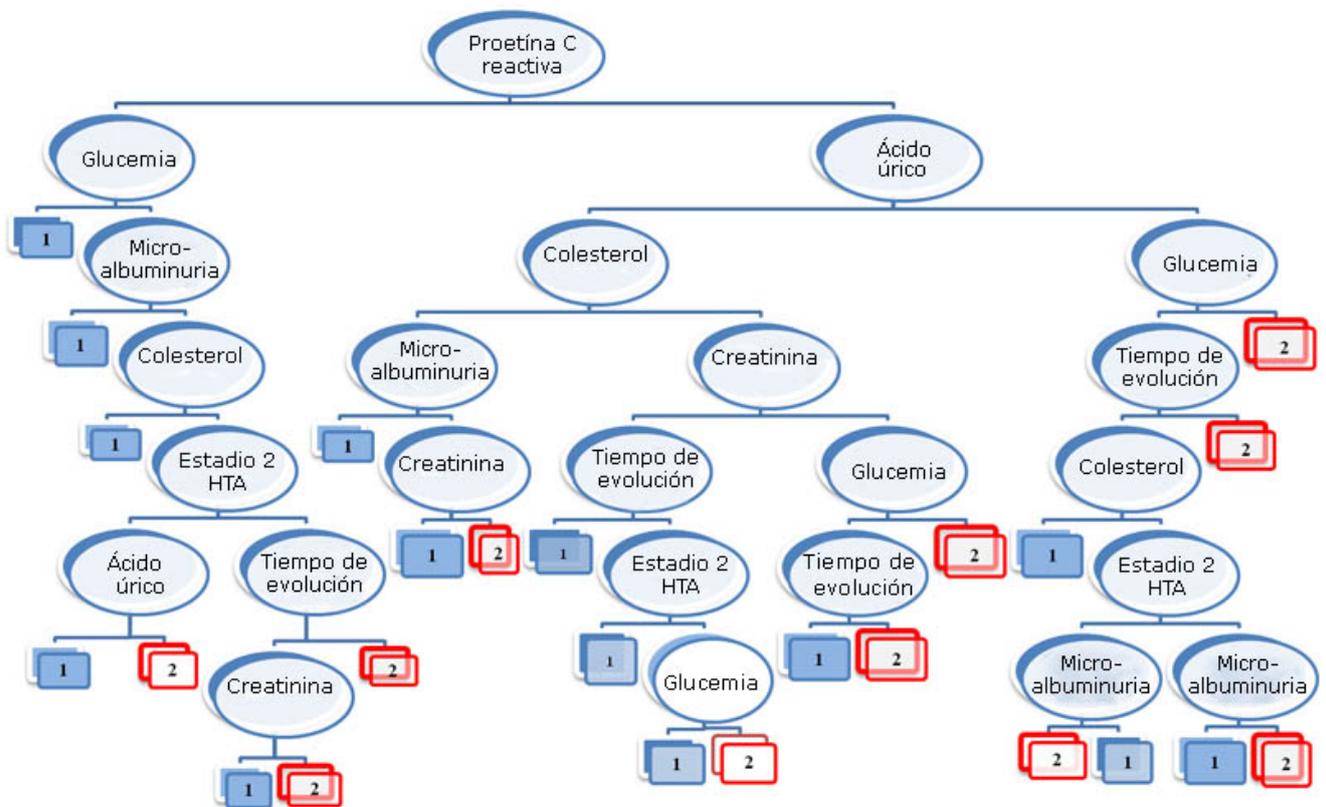
*p = 0,000 indica < 0,001.

N= 1 200.

Con el empleo del sistema que combina la técnica de KDD con DM se obtuvo un árbol de predicción para el desarrollo de la cardiopatía hipertensiva con adecuada capacidad predictiva (cuadro). Este proceso permitió predecir el riesgo de desarrollar la cardiopatía hipertensiva a un 82,598 % de los pacientes. Con un área bajo la curva ROC para ambas clases de 0,861; una tasa de verdaderos positivos de 0,733 y de 0,921 para las clase 1 y 2, respectivamente; además, el método mostró buena precisión a la hora de clasificar a los enfermos en: clase 1 (pacientes con cardiopatía hipertensiva): 0,764 y clase 2 (pacientes sin cardiopatía hipertensiva): 0,854. El factor más importante lo constituyó la PCR, al constituir el eslabón principal del árbol, le siguieron en orden de importancia, la glucemia, el ácido úrico, el colesterol, la microalbuminuria, el estadio 2 de la HTA y el tiempo de evolución de la HTA mayor de 20 años (Fig.).

El diseño también permitió crear 14 reglas para la clase 1 (casos) y 11 para la 2 (testigos) autovalidadas, como se muestra en el cuadro. Así, por ejemplo, si el individuo tiene PCR y glucemia por encima de los valores de corte, entonces tiene una probabilidad de 92 % de desarrollar la cardiopatía hipertensiva; sin embargo, si la PCR está por encima de 4,5 mg/L, mientras que los valores de la glucemia, la

microalbuminuria, el colesterol y el tiempo de evolución están por debajo del corte y el enfermo está clasificado en el estadio 1 de la HTA, entonces tiene 94 % de probabilidad de no desarrollar la cardiopatía hipertensiva.



1: presencia del factor de riesgo. 2: ausencia del factor de riesgo.

Fig. Árbol para predecir el desarrollo de la cardiopatía hipertensiva.

Los expertos consideraron que el contenido del árbol cumplía de forma moderada, y mucho, con las condiciones que se evaluaban. Sobresale que el 87,5 % refirió que el contenido era razonable y comprensible, derivable de datos factibles de obtener y que los componentes fueron claramente definidos (tabla 4).

Tabla 4. Validez de contenido y presentación del árbol de predicción, según opiniones de expertos

Árbol de predicción						
Contenido	Nada		Moderada		Mucho	
	No.	%	No.	%	No.	%
Razonable y comprensible	0	0	1	12,5	7	87,5
Sensible a variaciones del fenómeno que se mide	0	0	3	37,5	5	62,5
Suposiciones básicas justificables	0	0	2	25,0	6	75,0
Componentes claramente definidos	0	0	1	12,5	7	87,5
Derivable de datos factibles de obtener	0	0	1	12,5	7	87,5

DISCUSIÓN

La cardiopatía hipertensiva es la más frecuente, precoz e importante complicación encontrada en el paciente hipertenso y posee una elevada mortalidad,^{1-4,12,13} por lo que la identificación de sus factores de riesgo es útil para poder encausar acciones de prevención de la enfermedad.

Múltiples guías clínicas e investigaciones han abordado la determinación y evaluación de estos factores como predictores de complicaciones cardiovasculares, pero las características poblacionales (etnia, hábitos, costumbres e, incluso, genéticas) y las modificaciones de la propia enfermedad en su evolución, hacen variar su expresión en diferentes latitudes, hechos que indican continuar investigando sobre la temática.

En un principio, las guías sobre HTA se habían centrado en el valor de la presión arterial como único parámetro que determinaba la necesidad y el tipo de tratamiento a seguir. Este concepto ha ido evolucionando y actualmente en la conducta ante un paciente hipertenso se debe incluir la cuantificación del riesgo cardiovascular individual y el daño a órganos diana clínico y subclínico.^{1,8,11,22}

En aras de dar solución a la problemática, se crearon numerosos modelos e índices basados en los principales factores de riesgo de las ECV, que contribuyen a predecir su desarrollo y pronóstico.²³⁻²⁸

En la presente investigación se logró la construcción de un árbol de toma de decisiones útil para predecir el desarrollo de la cardiopatía hipertensiva, así como la generación de reglas predictivas autovalidadas, cuya interpretación es sencilla y de fácil aplicación, y permiten establecer predicciones de la enfermedad objeto de estudio, con exactitud adecuada. La elevada tasa de verdaderos positivos, para ambas clases, indica que es poco probable que se clasifiquen de alta probabilidad de desarrollar la cardiopatía hipertensiva aquellos pacientes que no enfermaran.

Aunque las herramientas propuestas no son el reflejo del valor exacto del riesgo, se consideran aceptables. Las variables y su posición en el árbol son perfectamente razonables y se apoyan en la importancia que tienen los diferentes factores de riesgos citados en diversas investigaciones.^{1,6,7,10,11,16,23,25}

Basado en lo antes expresado, es lógico encontrar que la PCR, la glucemia, el ácido úrico, el colesterol, la microalbuminuria, el estadio 2 de la HTA y el tiempo de evolución de la hipertensión, sean las variables principales en el árbol y a la vez generen las reglas con mejor probabilidad predictiva.

La principal ventaja del árbol de predicción propuesto y el conjunto de reglas derivado, es su simplicidad y fácil aplicación, incluye sólo 8 parámetros clínico-complementarios que suelen estar a disposición de los médicos que laboran tanto en atención primaria de salud como secundaria, en todo el territorio nacional, sin requerir exámenes complejos y costosos. Además, los *componentes del árbol se obtienen mediante un riguroso análisis, empleando el procedimiento de descubrimiento de conocimientos en bases de datos y minería de datos*, que permitió la creación de un algoritmo predictivo autovalidado. Tanto la sensibilidad como el valor predictivo positivo elevado, sugieren ventajas sobre las escalas de riesgos de la ECV conocidas, las cuales tienen entre sus principales limitaciones, la sensibilidad, la especificidad y el elevado coste.²⁹⁻³¹

La principal limitación de las herramientas propuestas, deriva del tamaño de la muestra examinada, si se tiene en cuenta que los modelos predictivos de las ECV se han validado en grandes muestras poblacionales. Hecho que indica la necesidad de comprobar la utilidad clínica en otras poblaciones de pacientes hipertensos, así como probar su efectividad y eficiencia.

Se concluye que en la presente investigación se derivó un árbol de toma de decisiones y un conjunto de reglas, que permitieron predecir el riesgo de desarrollar la cardiopatía hipertensiva en los individuos hipertensos, con alta sensibilidad y elevado valor predictivo positivo, por lo que su introducción en la práctica clínica, mejorará la evaluación del paciente con esta entidad.

REFERENCIAS BIBLIOGRÁFICAS

1. The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation and Treatment of High Blood Pressure. JAMA [Internet]. 2003 [citado 15 May 2004];289(19):2560-71. Disponible en: <http://jama.jamanetwork.com/article.aspx?articleid=185953>
2. Sociedad Española de Hipertensión-Liga Española para la Lucha contra la Hipertensión Arterial SEH-LELHA. Guía sobre el diagnóstico y el tratamiento de la hipertensión arterial en España 2005. Hipertensión [Internet]. 2005 [citado 12 Abr 2007];22(Supl 2):1-84. Disponible en: <http://www.seh.lelha.org/guia05.htm>
3. Diamond JA, Phillips RA. Hypertensive heart disease. Hypertens Res [Internet]. 2005 [citado 15 May 2007];28:191-202. Disponible en: <http://www.nature.com/hr/journal/v28/n3/pdf/hr200525a.pdf>
4. Mitchell JA, Ventura HO, Mehra MR. Early recognition and treatment of hypertensive heart disease. Curr Opin Cardiol. 2005;20:282-9.

5. López B, González A, Querejeta R, Díez J. The use of collagenderived serum peptides for the clinical assessment of hypertensive heart disease. *J Hypertens*. 2005;23:1445-51.
 6. Díez J, Frohlich ED. A translational approach to hypertensive heart disease. *Hypertension* [Internet]. 2010 [citado 29 Feb 2010];55:1-8. Disponible en: <http://hyper.ahajournals.org/content/55/1/1>
 7. Egan BM, Zhao Y, Axon RN. US trends in prevalence, awareness, treatment, and control of hypertension, 1988-2008. *JAMA* [Internet]. 2010 [citado 21 Ene 2011];303(20):2043-50. Disponible en: <http://jama.jamanetwork.com/article.aspx?articleid=185953>
 8. Mancia G, Laurent S, Agabiti-Rosei E, Ambrosioni E, Burniere M, Caulfield MJ, et al. Reappraisal of European guidelines on hypertension management: a European Society of Hypertension Task Force document. *J Hypertens* [Internet]. 2009 [citado 30 Feb 2010];27(11):1-38. Disponible en: <http://www.ish.org.il/2009GuidelinesESH.pdf>
 9. Tousoulis D, Koumallos N, Antoniadis C, Antonopoulos AS, Bakogiannis B, Milliou A, et al. Genetic polymorphism on type 2 receptor of angiotensin II, modifies cardiovascular risk and systemic inflammation in hypertensive males. *Am J Hypertens* [Internet]. 2010 [citado 14 Dic 2010];23:237-42. Disponible en: <http://ajh.oxfordjournals.org/content/23/3/237.full>
 10. García Barreto D, Álvarez González J, García Fernández R, Valiente Mustelier J, Hernández Cañero A. La hipertensión arterial en la tercera edad. *Rev Cubana Med* [Internet]. 2009 Jun [citado 20 Ago 2010];48(2). Disponible en: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0034-75232009000200007&lng=es&nrm=iso&tlng=es
 11. Mancia G, De Backer G, Dominiczak A, Cifkova R, Fagard R, Germano G, et al. Guidelines for the Management of Arterial Hypertension: The Task Force for the Management of Arterial Hypertension of the European Society of Hypertension (ESH) and of the European Society of Cardiology (ESC). *J Hypertens* [Internet]. 2007 [citado 16 May 2008];25:1105-87. Disponible en: <http://www.escardio.org/guidelines-surveys/esc-guidelines/GuidelinesDocuments/guidelines-AH-FT.pdf>
 12. Devereux RB, Reichek N. Echocardiographic determination of left ventricular mass in man. Anatomic validation of the method. *Circulation* [Internet]. 1977 [citado 13 Ene 2004];55:613-8. Disponible en: <http://circ.ahajournals.org/content/55/4/613.long>
 13. Graf K, Schaefer-Graf UM. Is Smad3 the key to inflammation and fibrosis in hypertensive heart disease? *Hypertension* [Internet]. 2010 [citado 28 Ene 2011];55:1088-9. Disponible en: <http://hyper.ahajournals.org/content/55/5/1088.long>
 14. Khouri MG, Peshock RM, Ayers CR, de Lemos JA, Drazner MH. A 4-tiered classification of left ventricular hypertrophy based on left ventricular geometry: the Dallas Heart Study. *Circ Cardiovasc Imaging* [Internet]. 2010 [citado 28 Nov 2010];3:164-71. Disponible en: <http://circimaging.ahajournals.org/content/3/2/164.full.pdf>
 15. Beaumont J, Arias T, López B, González A, Ravassa S, Hermida N, et al. Avances en cardiopatía hipertensiva. Mecanismos de remodelado implicados en la transición de la hipertrofia a la insuficiencia cardíaca. *Rev Esp Cardiol* [Internet]. 2007 [citado 31 Mar 2009];7(Supl):14F-21F. Disponible en: <http://www.revespcardiol.org/es/avances-cardiopatia-hipertensiva-mecanismos-remodelado/articulo/13110828/>
-

16. Drazner MH. The Progression of Hypertensive Heart Disease. *Circulation* [Internet]. 2011 [citado 16 May 2011];123:327-34.
Disponible en: <http://circ.ahajournals.org/content/123/3/327.long>
17. Fuster V, Rydén LE, Cannom DS, Crijs HJ, Curtis AB, Ellenbogen KA, et al. ACCF/AHA/HRS focused updates incorporated into the ACC/AHA/ESC 2006 guidelines for the management of patients with atrial fibrillation. *J Am Coll Cardiol* [Internet]. 2011 [citado 15 Mar 2011];123:104-23.
Disponible en: <http://circ.ahajournals.org/content/123/1/104.full>
18. Pepys MB, Hirschfield GM. C-reactive protein: a critical update. *J Clin Invest* [Internet]. 2003;111(12):1805-12. [citado 15 Mar 2004] Disponible en: <http://www.jci.org/articles/view/18921> doi 10.1172/JCI18921
19. MICROALBUMIN Latex L. Quantitative turbidimetric latex assay for the measurement of Microalbumin in human urine. [Internet]. 1996 [citado 12 May 2003].
Disponible en:
http://www.gdsrl.com/upload/prodotti/gd8439_00_microalbumin_1.pdf
20. Reyes Saldaña JF, García Flores R. Descubrimiento de conocimientos en bases de datos. *Ingenierías* [Internet]. 2005 [citado 25 Ene 2009];8(26):37-47.
Disponible en:http://ocw.univalle.edu.co/ocw/ingenieria-de-sistemas-telematica-y-afines/descubrimiento-de-conocimiento-en-bases-de-datos/Course_listing
21. Alonso C. WEKA: Waitako Environment for Knowledge Analysis. Introducción básica. Departamento de Informática Universidad de Valladolid. [Internet]. 2006 [citado 27 Ago 2009]. Disponible en:
<http://www.infor.uva.es/~calonso/IAII/Aprendizaje/weka/IntroduccionWeka.pdf>
22. Pérez Caballero MD, Dueñas Herrera A, Alfonso Guerra JP, Vázquez Vigoa A, Navarro Despaigne D, Hernández Cueto M, et al. Hipertensión arterial. Guía para la prevención, diagnóstico y tratamiento. Comisión Nacional Técnica Asesora del Programa de Hipertensión Arterial. La Habana: Editorial Ciencias Médicas; 2008.
23. D'Agostino RB, Grundy S, Sullivan LM, Wilson P, for the CHD Risk Prediction Group. Validation of the Framingham Coronary Heart Disease Prediction Scores. Results of a multiple ethnic groups investigation. *JAMA* [Internet]. 2001 [citado 14 Ene 2007];286:180-7. Disponible en:
<http://jama.jamanetwork.com/article.aspx?articleid=193997>
24. Marrugat J, Solanas P, D'Agostino R, Sullivan L, Ordovás J, Cerdán F, et al. Estimación del riesgo coronario en España mediante la ecuación de Framingham calibrada. *Rev Esp Cardiol* [Internet]. 2003 [citado 14 Ene 2007];56(3):253-61.
Disponible en: <http://www.revespcardiol.org/es/estimacion-del-riesgo-coronario-espana/articulo/13043951/>
25. Marshall T. The effect of blood pressure and cholesterol variability on the precision of Framingham cardiovascular risk estimation: a simulation study. *J Hum Hypertens* [Internet]. 2010 [citado 18 Ene 2011];24:1-8. Disponible en:
<http://www.nature.com/jhh/journal/v24/n10/full/jhh2009114a.html>
26. Woodward M, Brindle P, Tunstall-Pedoe H. Adding social deprivation and family history to the cardiovascular risk assessment: the ASSIGN score from the Scottish Heart Health Extended Cohort (SHHEC). *Heart* [Internet]. 2007 [citado 26 Dic 2007];2:172-6. Disponible en: <http://heart.bmj.com/content/93/2/172.long>
27. Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, May M, Brindle P. Derivation and validation of QRISK, a new cardiovascular disease risk score for the United Kingdom: prospective open cohort study. *BMJ* [Internet]. 2007 [citado

26 Dic 2007];335:136-41.

Disponible en: <http://www.bmj.com/content/335/7611/136>

28. D'Agostino RB, Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, et al. General cardiovascular risk profile for use in primary care the Framingham heart study. *Circulation* [Internet]. 2008 [citado 26 Jun 2009];117:743-53. Disponible en: <http://circ.ahajournals.org/content/117/6/743.long>

29. Baena-Díez JM, Ramos R, Marrug J. Capacidad predictiva de las funciones de riesgo cardiovascular: limitaciones y oportunidades. *Rev Esp Cardiol* [Internet]. 2009 [citado 25 Ago 2010];9(Supl):4B-13B. Disponible en: <http://www.revespcardiol.org/es/capacidad-predictiva-las-funciones-riesgo/articulo/13137261/>

30. Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, May M, Brindle P. Derivation and validation of QRISK, a new cardiovascular disease risk score for the United Kingdom: prospective open cohort study. *BMJ* [Internet]. 2007 [citado 26 Dic 2011];335:136-41.

Disponible en: <http://www.bmj.com/content/335/7611/136>

31. Buitrago F, Cañón-Barroso L, Díaz-Herrera N, Cruces-Muro E, Escobar-Fernández M, Serrano-Arias JM. Comparación de las tablas REGICOR y SCORE para la clasificación del riesgo cardiovascular y la identificación de pacientes candidatos a tratamiento hipolipemiante o antihipertensivo. *Rev Esp Cardiol*. [Internet]. 2007 [citado 28 Feb 2009];60:139-47. Disponible en:

<http://www.revespcardiol.org/es/comparacion-las-tablas-regicor-score/articulo/13099460/>

Recibido:20 de noviembre de 2013.

Aceptado: 13 de mayo de 2014.

Dr. *Alexis Álvarez Aliaga*. Hospital General Universitario "Carlos Manuel de Céspedes". Carretera Central km 1 vía a Santiago de Cuba. Bayamo, Granma, Cuba. alexis.grm@infomed.sld.cu