

## SIMILITUD DE MARCADORES PREDICTIVOS DEL DAÑO CEREBRAL EN HIPERTENSOS ASINTOMÁTICOS

Carlos N. Bouza Herrera<sup>1\*</sup>, Jose F. Garcia<sup>2\*\*</sup>, Anaí Garcia<sup>3\*\*\*</sup>, Carmelo Herrera<sup>4\*\*\*\*</sup>, Teresa La Rosa Chibas<sup>5\*\*\*\*\*</sup> y Pedro A. Valdés<sup>6 \*\*\*\*\*</sup>

\*Universidad de La Habana, Cuba

\*\*Universidad Juárez Autónoma de Tabasco, México

\*\*\*Escuela Nacional de Salud Pública Cuba

\*\*\*\*Universidad Las Palmas Gran Canaria, España

\*\*\*\*\*Centro Nacional de Neurociencias, Cuba

### ABSTRACT

The similarity among predictive markers is studied. Similarity measures are used. They evaluate the affectivity of the markers without assuming hypothesis on the distribution of the statistics. These hypothesis are not necessary true. The results are similar to those obtained by using the technique proposed by La Rosa (2011) in her MSC. Thesis. Now some dubiously effective markers are eliminated.

**KEY WORDS:** curve ROC, standard ordinal, similarity measures

### RESUMEN

Se hace un estudio de la similitud presentada entre marcadores predictivos del daño cerebral. Se utilizan medidas de similitud. Estas dan una idea de la efectividad de los marcadores que no dependen de hipótesis sobre la distribución de estadísticos, lo cual no necesariamente son válidas. Los resultados son parecidos aunque se discriminan algunos de los considerados dudosamente efectivos usando la técnica propuesta por La Rosa (2011) en su tesis de maestría

**PALABRAS CLAVE** curva ROC, estándar ordinal, medidas de similaridad

## 1. INTRODUCCIÓN

La exactitud de una prueba diagnóstica es evaluada comúnmente en función de su sensibilidad y la especificidad. El análisis de la Curva de Características Operacionales del Receptor, más conocida por la sigla ROC (Receiver Operating Characteristic) es muy utilizada con este fin. Esta metodología se desarrolló dentro de la Teoría de la Decisión, aplicándose en la resolución de problemas de detección de señales. Además se utilizan en psicología para medir la eficacia de una droga nueva o de un método de diagnóstico contra los ya existentes, en epidemiología y en otras investigaciones médicas, ver La Rosa (2011).

La Hipertensión Arterial (HTA) es de importante prevalencia afectando la salud de más de un billón de personas en el mundo. Es conocida su capacidad para reducir sustancialmente la duración y calidad de vida de los pacientes generando complicaciones cardíacas, renales y del sistema nervioso central y es un factor de riesgo y de aceleración de la aterosclerosis. En Cuba se reportaron en el 2009 cifras del 33,9% en adultos mayores de 18 años y en el 2010, El Grupo Nacional de HTA del estudio de escolares en Villa Clara reportó en el año 2010 la prevalencia de 27 %.

<sup>1</sup> [bouza@matcom.uh.cu](mailto:bouza@matcom.uh.cu)

<sup>2</sup> [ifer55@hotmail.com](mailto:ifer55@hotmail.com)

<sup>3</sup> [anaigf@yahoo.es](mailto:anaigf@yahoo.es)

<sup>4</sup> [vihega@movistar.es](mailto:vihega@movistar.es)

<sup>5</sup> [teresa.lr@cneuro.edu.cu](mailto:teresa.lr@cneuro.edu.cu)

<sup>6</sup> [peter@cneuro.edu.cu](mailto:peter@cneuro.edu.cu)

Varios estudios han demostrado una asociación independiente de las cifras de HTA con la demencia y el deterioro cognitivo. Hallazgos más recientes, destacan el papel de la HTA en la patogenia de la Enfermedad de Alzheimer y el Deterioro Cognitivo Leve.

Entre las pruebas que miden el funcionamiento del cerebro, se encuentra la Resonancia Magnética (RM), sin embargo pocos estudios utilizan como marcadores de daño cerebral las lesiones silentes debido al elevado costo y la baja disponibilidad de las técnicas de Resonancia Magnética, que imposibilita su uso masivo en la evaluación del daño sub-clínico cerebral en la HTA.

El costo de adquisición de un equipo de RMN se ubica en el rango de los millones de dólares en dependencia de la potencia (tesla), siendo mayor cuanto mayor es esta. Varios estudios previos han señalado a la RMN como técnica de diagnóstico no costo efectiva o sujeta a condiciones muy estrictas para que su costo se justifique. Ejemplo de ello es el estudio en el que se evaluaron dos estrategias de pesquisas poblacionales para detectar tumores en pacientes con psicosis y se encontró que no era costo efectivo el empleo de la RMN, pues el costo efectividad superó las 20000 libras esterlinas por año de vida ajustado por calidad. (Albon, 2008) valor umbral establecido por las autoridades del reino unido para definir si es factible una tecnología sanitaria para un determinado uso. Otro caso es el estudio de la RMN para detectar cáncer de próstata, realizado por Stadlbauer en Alemania que arrojó un incremento del costo no justificado por los resultados clínicos (Stadlbauer, 2011). En el campo de las enfermedades neurodegenerativas el estudio realizado por Biasutti para detección precoz de alzheimer reportó que sólo se consideraría la RMN como estrategia costo efectiva si se lograran elevados niveles de especificidad tales que permitiera un diagnóstico en estadios tempranos. (Biasutti, 2012) El Proyecto “Marcadores precoces de lesión cerebral en pacientes hipertensos neurológicamente asintomáticos del municipio La Lisa”, que dirige el Centro de Neurociencias de Cuba, tiene dentro de los objetivos de su investigación la determinación de variables (pruebas) fisiológicas que miden el funcionamiento del cerebro. Son pruebas sencillas y económicamente viables, que necesitan evaluarse, para determinar su precisión y validez diagnóstica que permita valorar la posibilidad de constituir biomarcadores, los cuales pueden ser utilizados en estudios poblacionales, en aras de implementar estrategias de intervención terapéutica. En el estudio de pacientes hipertensos se tuvo en cuenta que el óptimo funcionamiento del cerebro depende de muchas variables que inciden directamente en la actividad del mismo. Algunas de esas variables constituyen biomarcadores que son los cambios medibles para detectar los estados iniciales e intermedio de un proceso patológico que permite la toma de decisiones apropiadas. Bouza et al (2011) realizaron un estudio de un grupo de biomarcadores y estableció sus potencialidades para detectar daños cerebrales. En este trabajo vamos a estudiar la similitud de estos usando diversas medidas.

## 2. LAS CURVAS ROC.

Consideremos la razón de verosimilitud  $L_{sr} = \frac{P(S|S)}{P(S|R)}$ ;  $P(S|S)$  es la probabilidad de la sensación dada la señal y  $P(S|R)$  es la probabilidad de la sensación dado el ruido solo. En medicina se tienen que la Señal es si el paciente esta enfermo o no y la Sensación es el diagnóstico que el médico hace. El problema del diagnóstico es modelado al fijar un valor umbral  $c$  decidir que el paciente esta enfermo si  $L_{sr} < C$ . A partir de  $L_{sr}$  se definen los índices de precisión y exactitud diagnóstica.

Existen clasificadores que, de forma natural producen la probabilidad (puntuación o score) que mide el grado con que este es miembro de la clase de interés.

Las pruebas de diagnóstico que producen puntuaciones o rangos pueden ser usadas con un punto de corte. Utilizando la notación de una prueba continúa transformada a binaria definida por, ver Pepe (1998): Positivo si  $Y \geq c$ , Negativo si  $Y < c$ , De ahí las definiciones siguientes:

**Definición 2.1.** . Dado el rango particionado en subintervalos limitados por los valores  $x_0, x_1, x_2, \dots, x_c$  y fijando uno de ellos como punto de umbral ( $c$ ), tomando  $x_{VN}$  como la puntuación para un caso verdaderamente

negativo, ( $x = V_{N(xc)}$ ), es la proporción de verdaderos negativos, (los que están por debajo del punto de corte en la prueba y realmente son negativos). □

**Definición 2.2.** La función  $R$ , definida en el intervalo  $[0,1]$  en que  $R = 1 - S^- = S^+$  recibe el nombre de función ROC □

Esta igualdad genera una curva que se mueve de  $(0,0)$  a  $(1,1)$  llamado curva ROC (Chap 2006). □

La curva ROC igualmente se puede definir como el conjunto de posibles fracciones de verdaderos positivos y falsos positivos obtenidas por la utilización de diferentes puntos de corte:

$$ROC(c) = \left( \frac{FP(c)}{FVP(c)}, FVP(c) \right), c \in (-\infty, \infty)$$

Como se había mencionado anteriormente, Existe una amplia literatura sobre los métodos estadísticos paramétricos, no paramétricos, y semiparamétricos para estimar y comparar las curvas ROC y sus índices asociados, cuando el estándar de oro es a escala binaria (Braga, 2000).

El método no paramétrico por su flexibilidad es muy utilizado para la estimación de la ROC con estándar de oro binario y no binario. Entre los métodos no paramétricos se encuentra la estimación empírica de la ROC.

Un método aplicable a datos continuos fue desarrollado por Metz et al (1980). Este método es llamado semiparamétrico.

**Definición 2.3.** El método de Metz et al. es implementado por el algoritmo siguiente

Paso 1: Agrupar los datos en categorías ordenadas

Paso 2: Aplicar un algoritmo paramétrico para crear una curva ROC suave. □

Este método puede ser implementado usando el software libremente disponible en <http://www-radiology.uchicago.edu/krl/toppage11.htm>.

Es conocido que  $\lim_{t \rightarrow 0} ROC(t) = 0$  y  $\lim_{t \rightarrow 1} ROC(t) = 1$ . Si se denotan por  $F$  y  $G$  las funciones de distribución del resultado de un clasificador  $Y$  en las poblaciones enferma y no enferma, la función ROC se puede escribir como:  $ROC = 1 - F(G^{-1}(1-t))$ ,  $t \in (0,1)$ . La curva ROC queda así definida en el intervalo  $(0,1)$  como  $ROC(0) = 0$  y  $ROC(1) = 1$  con pendiente

$$\frac{\partial ROC(t)}{\partial t} = \frac{f(F^{-1}(t))}{g(F^{-1}(t))}$$

donde  $f$  y  $g$  son las funciones de densidad de los resultados de la variable  $Y$ , ver Pepe (2003) y Krzanowski y Hand (2009).

### 3. INDICADORES PARA LA EVALUACIÓN DE LAS ROC

Entre los indicadores de la ROC que son utilizados en la evaluación de las pruebas diagnósticas se encuentran:

- El área bajo la curva ROC, (AUC, Area Under the Curve): Área entre la curva ROC y la línea de no-discriminación.
- Índice de sensibilidad o  $d'$ : La distancia entre la media de la distribución de los verdaderos positivos y los verdaderos negativos dividido por su desviación típica, bajo el supuesto de que ambas distribuciones son normales con la misma desviación típica.

c. El índice de Youden:  $J = S^+ + S^- - 1$ ,  $J \in [0, 1]$ . Se alcanza 1, para un perfecto ensayo y 0, si no tiene valor diagnóstico.

Existen trabajos que utilizan este índice como parámetro discriminatorio entre los dos errores (Bouza 2006).

### 3.1. Método de estimación del área bajo la curva ROC.

Si se utiliza solo la información obtenida empíricamente se debe tener en cuenta la definición siguiente:

**Definición 3.1.** Sean  $X$  e  $Y$  dos puntajes que se obtienen al tomar de manera aleatoria un sujeto negativo (sano) y (enfermo) respectivamente, entonces,  $\theta = \Pr\{X_{VP} > X_{VN}\}$ ,  $\theta$ , mide el área bajo la curva ROC. □

Los valores posibles de AUC se extienden a partir de la 0.5 (ninguna capacidad de diagnóstico) a 1.0 (capacidad de diagnóstico perfecta).

En la práctica la evaluación de un test se lleva a cabo analíticamente usando la regla

$$A = \begin{cases} \text{excelente} & \text{si } x \geq 0.9 \\ \text{Igualmente bueno} & \text{si } x \in [0.8, 0.9] \\ \text{Es aceptable} & \text{si } x \in [0.6, 0.8] \\ \text{Es deficiente} & \text{si } x \in [0.5, 0.6] \\ \text{Es malo} & \text{si } x \leq 0.5 \end{cases}$$

Existen diversas fórmulas propuestas por diferentes autores (DeLong 1988), (Zweig 1993) para estimar el error estándar de esta área. Se utiliza usualmente:

$$ES_{\theta} = \sqrt{\frac{\theta(1-\theta) + (Q_A - 1)(Q_1 - \theta^2) + (Q_N - 1)(Q_2 - \theta^2)}{n_A n_N}}$$

$Q_1$  = probabilidad que los sujetos aleatoriamente anormales tendrán más resultados anormales que un sujeto normal elegido aleatoriamente.

$Q_2$  = probabilidad que los sujetos aleatoriamente anormales tendrán más resultados anormales que dos sujetos normales elegidos aleatoriamente.

$n_A$  = número de sujetos anormales.

$n_N$  = número de sujetos normales.

Para el estudio se toman muestras dentro de una población de casos verdaderamente negativos y de otra de verdaderamente positivos, cuyas distribuciones se suponen gaussianas. Estas se solapan para un umbral o punto de corte, determinado por la experiencia del observador y la información previa que se tenga sobre la patología del paciente.

Cuando la curva ROC se genera por el método empírico el área puede calcularse mediante la regla trapezoidal, es decir, como la suma de las áreas de todos los rectángulos y trapecios (correspondientes a los empates) que se pueden formar bajo la curva.

La hipótesis nula es que  $X$  no es un discriminador útil es  $H_0 : \theta = \Pr\{X_{VP} > X_{VN}\} = 0.5$ .

El procedimiento de Wilcoxon, (Alan, 2001) se basa en hacer todas las posibles comparaciones entre los valores  $X_{VP(ij)}$  de  $n_{VP(i)}$  contra los valores  $X_{VN(ij)}$  de  $n_{VN(j)}$ , de acuerdo con la regla de decisión (Zhou 2003) :

$$T(X_p, X_N) = \begin{cases} 1 & \text{si } X_{VP} > X_{VN} \\ \frac{1}{2} & \text{si } X_{VP} = X_{VN} \\ 0 & \text{si } X_{VP} < X_{VN} \end{cases}$$

Promediando los casos (sujetos) a lo largo de todas las comparaciones  $n_{VP} \times n_{VN}$ , se calcula:

$$\theta = W = \frac{1}{n_{VP} n_{VN}} \sum_1^{n_E} \sum_1^{n_N} T(X_{VP}, X_{VN})$$

Este depende del ordenamiento y resume el área bajo de la curva.

### 3.2 Comparación de dos curvas ROC.

Al comparar los resultados de un clasificador a evaluar y el patrón de oro, existen cuatro posibilidades que pueden resumirse en una tabla de contingencia de 2 x 2 que se ilustran a continuación (tabla1).

	RESULTADO DEL BIOMARCADOR ESTUDIADO	DIAGNÓSTICO VERDADERO (PRUEBA DE REFERENCIA)	
		Enfermo	Sano
PRUEBA ESTUDIADA	Positivo	Verdadero positivo (VP)	Falso positivo (FP)
	Negativo	Falso negativo (FN)	Verdadero negativo (VN)

**Tabla 1.** Comparación de los resultados de un clasificador a evaluar con el diagnóstico verdadero.

A partir de la tabla, se pueden obtener los estimadores que se relacionan a continuación:

$$S^+ = \frac{VP}{VP + FN} = FVP \quad \text{y} \quad S^- = \frac{VN}{VN + FP} = FVN = 1 - FFP$$

Siendo FVN fracción de los verdaderos negativos, FFP fracción de los falsos positivos  $S^+$  sensibilidad y  $S^-$  especificidad. La razón de probabilidad para un resultado positivo ( $L_{R+}$ ) indica cuánto aumenta las probabilidades de la enfermedad cuando un resultado es positivo:

$$L_{R+} = S^+ / (1 - S^-)$$

También se define la razón de probabilidad de un resultado negativo como:

$$L_{R-} = (1 - S^+) / S^-$$

La razón de probabilidad para un resultado negativo ( $L_{R-}$ ) indica cuánto disminuyen las probabilidades de la enfermedad cuando un resultado es negativo.

Si se desea evaluar la efectividad de marcador cuyos resultados quedan determinados por una variable continua o al menos ordinal, la generalización a estas situaciones se consiguen mediante la elección de distintos niveles de decisión o valores de corte que permitan una clasificación dicotómica de los valores de la prueba según sean superiores o inferiores al valor elegido. Se divide el rango de valores en subintervalos  $I_0, I_1, I_2, \dots, I_m$  y se contabilizan cuántos pacientes de la muestra obtuvieron valores de la prueba dentro de cada intervalo. Fijando uno de ellos como punto de corte (cutoff) o umbral discriminante, se crea una tabla  $2 \times 2$ . Suponiendo que el test es positivo si  $Z \leq I_k$ , y negativo si  $Z > I_k$ , se realiza el mismo análisis anteriormente expuesto y se obtiene la tabla siguiente llamada matriz de confusión. Esta matriz es la base para el cálculo de varios indicadores comúnmente utilizados en el sistema de diagnóstico.

Resultados del biomarcador estudiado	Diagnóstico verdadero (Prueba de referencia)	
	Enfermo	Sano
$P: Z \leq I_k$	$VP = x_{01} + x_{11} + \dots + x_{k1}$	$FP = x_{00} + x_{10} + \dots + x_{k0}$
$E: Z > I_k$	$FN = x_{k+1,1} + x_{k+2,1} + \dots + x_{m+1,1}$	$VN = x_{k+1,0} + x_{k+2,0} + \dots + x_{m+1,0}$

**Tabla 2.** Posibles resultados de una marcador para diferente puntos de corte.

Tomamos  $x_{ij}$  como el número de pacientes enfermos (si  $j=1$ ) o sanos (si  $j=0$ ) a quienes la prueba diagnóstica ha dado un valor en el intervalo  $(I_{i-1}, I_i]$ . Cuando  $i=0$  ó  $i=m+1$ , los intervalos son semirrectas. La diferencia esencial con el caso más simple es que se puede contar no con un único par de valores de sensibilidad y especificidad que definan la exactitud de marcadores, sino más bien con un conjunto de pares correspondientes cada uno a un distinto nivel de decisión. La posición relativa de las líneas indica la precisión relativa de las pruebas. Por ejemplo, una curva situada por encima y a la izquierda de otra indica una mayor precisión.

Desde un punto de vista estadístico el problema que se enfrenta es valorar si la diferencia observada entre las áreas bajo las curvas calculadas para dos pruebas distintas es debida a la variabilidad inherente al muestreo, o si es atribuible a una diferencia real en la exactitud de ambas pruebas. Se puede enunciar entonces este problema como un contraste de la hipótesis nula de igualdad de las dos áreas, que denotaremos por  $\theta_1, \theta_2$ , frente a una alternativa bilateral.

En general se sigue el siguiente procedimiento:

Paso 1 Se escoge la hipótesis nula que va a estar relacionada con los parámetros de la curva ROC y se plantea el problema de prueba de hipótesis

$$H_0: \theta_1 - \theta_2 = 0 \quad \text{vs} \quad H_1: \theta_1 - \theta_2 \neq 0$$

Paso 2 Estimar los parámetros relevantes de las dos curvas

Paso 3 Escoger el test estadístico que permita que la hipótesis nula sea verdadera.

La prueba que se desarrolla usualmente utiliza como estadístico del contraste (Hanley 1983)

$$Z = \frac{\theta_1 - \theta_2}{\sqrt{EE_{\theta_1}^2 - EE_{\theta_2}^2 - 2REE_{\theta_1}EE_{\theta_2}}} \sim N(0,1)$$

Siendo  $\theta_1, \theta_2$  las áreas observadas  $EE_{\theta_1}$  y  $EE_{\theta_2}$  sus errores estándar respectivamente del ABC y  $R$  la correlación entre  $\theta_1, \theta_2$ . Para un nivel de significación  $\alpha$  se rechaza la hipótesis nula cuando  $|z| > Z_{\alpha/2}$ , siendo

$Z_{\alpha/2}$  el cuantil de orden  $1 - \alpha/2$  de una distribución normal estándar.

La paquetería comercial presenta procedimientos para cuando la prueba de referencia es una variable en escala binaria, lo que constituye una limitación para el estudio que se desea obtener. En Bouza et al (2011) se presenta un software en el que se utiliza el software SPSS para aplicar la metodología de la ROC de estándar binario a través del programa en R Nonbin ROC. Se utilizó la propuesta de Obuchowski (2005) para el cálculo de la exactitud de una prueba de diagnóstico cuando el estándar de oro es medido en una escala continua, ordinal o nominal.

## **4. LOS BIOMARCADORES**

### **4.1. El problema médico tratado**

La hipertensión arterial no tratada o no controlada afecta la estructura y función de varios órganos y sistemas a los que se les denomina “órganos diana” e incluyen el sistema nervioso central, arterias periféricas, corazón y riñones, principalmente. La asociación entre la presión arterial y el riesgo de cardiopatías, infarto agudo de miocardio, derrame cerebral y enfermedades renales es muy elevada e independiente de otros factores de riesgo. Tanto el grado de elevación de la tensión arterial sistólica como diastólica se correlacionan con el riesgo de desarrollar complicaciones vasculares, estableciéndose una relación continua, consistente e independiente de otros factores de riesgo, lo cual significa que a medida que aumenta la presión arterial también aumenta el riesgo, siendo esta situación individual para cada sujeto.

En el cerebro la HTA causa lesión en grandes vasos que se traduce clínicamente en Enfermedad Cerebrovascular Aguda (isquémica y hemorrágica). En Cuba las tasas de mortalidad por Enfermedad Cerebrovascular se han incrementado gradualmente en el último decenio y es actualmente la 3ra causa de muerte, donde luego de un decrecimiento en el 2002 (68.3 por 100 mil habitantes.) se incrementa nuevamente en el 2005 (78.1) (Miranda, 2004), lo cual constituye una alerta, unido al incremento de la prevalencia de hipertensos en el país (30.6 % en zonas urbanas con solo un 26 a 35 % de pacientes controlados) (MINSAP) alcanzando un tasa de prevalencia total de 194,5 con bajo control de la HTA en los diferentes territorios. La Habana es la provincia que encabeza la lista de fallecidos por esta causa, con una mortalidad en el 2008 (110,5 por 100 000 habitantes.) que supera ampliamente a la media nacional (82,3 por 100 000 hab.) (Situación de Salud en Cuba, 2008).

La HTA, causa además lesión de pequeños vasos cerebrales expresándose clínicamente con deterioro cognitivo vascular, demencia vascular y Enfermedad de Alzheimer, lo cual ha sido reportado en varios estudios (Obisesan, 2009). Las acciones dirigidas a prevenir dicho deterioro deben instaurarse precozmente, antes del desarrollo de una demencia establecida (Atienza, 2000).

Las Guías Europeas para el manejo de la HTA (Mancia, 2007) establecen el abordaje y la intervención y tratamiento a la HTA según la estratificación del riesgo del paciente y el daño a órganos diana, estando claramente definidos los marcadores de daño en riñón, corazón y sistema vascular periférico; pero sigue siendo una limitación la no existencia de marcadores precoces de fácil acceso para determinar el daño subclínico en el cerebro, a pesar de que existen estudios (Henkens, 2009) que demuestran que en un elevado porcentaje de casos el daño cerebral precede al cardíaco y/o renal en los pacientes hipertensos. Estos estudios utilizan como marcadores las lesiones cerebrales silentes en Resonancia Magnética que presentan los pacientes, pero los estudios de Imágenes de RM resultan muy costosos y poco disponibles. Se hace necesaria la determinación de Biomarcadores para la detección precoz de lesiones cerebrales en pacientes hipertensos.

Los marcadores biológicos o biomarcadores son los cambios medibles, ya sean estos bioquímicos, fisiológicos o morfológicos, que se asocian a la exposición, las consecuencias biológicas de la exposición es detectar los estados iniciales e intermedio de un proceso patológico, identificar a los individuos sensibles de una población, fundamentando la decisión de un diagnóstico certero que permita la tomar decisiones apropiadas.

Se realizó un estudio de los 132 pacientes de 21 a 71 años con diagnóstico clínico de hipertensión arterial, excluidos por tal motivo del estudio de Mapeo Cerebral Humano, pertenecientes al Policlínico Aleida Fernández en el Municipio la Lis, La Habana. De estos, por diversos motivos (cambio de domicilio, presencia de otras enfermedades crónicas concomitantes, presencia de dispositivos metálicos permanentes que contraindican la

realización de IRM, no voluntariedad, etc.) se constituyó una muestra final de 50 pacientes. De ellos 26 mujeres y 24 hombres. Con una edad promedio de 44 años, con un mínimo de 21 años y máximo de 65 años, con una desviación estándar de 31,11. El tiempo de evolución de la HTA oscilaba desde 3 meses a 30 años, con un tiempo promedio de 11 años de evolución de la enfermedad. En ellos se evaluaron biomarcadores que fueron medidos usando pruebas no invasivas y baratas. Un estudio de RM les fue aplicado para destormar su estado cerebral.

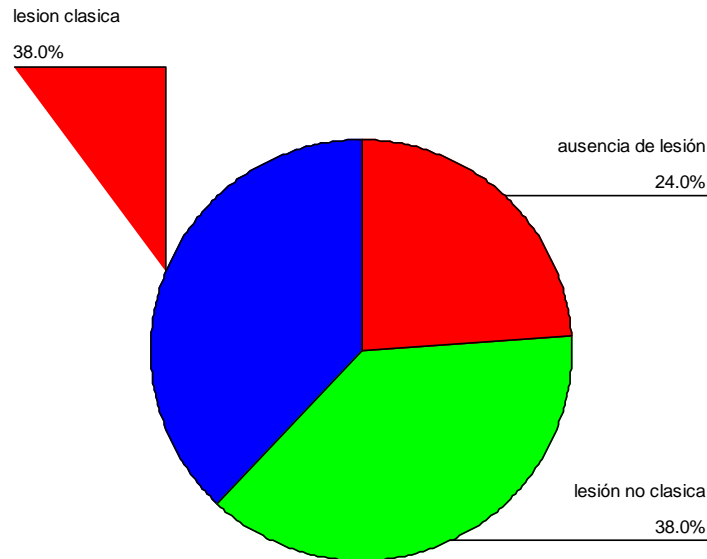
Las evaluaciones y exámenes complementarios a que fueron sometidos los pacientes se relacionan en La Rosa (2011).

#### 4.2 Biomarcadores vs Resonancia Magnética

Para la aplicación del modelo de la ROC no binaria se utiliza la función OrdROC, la cual forma parte de las funciones del software NonbiROC y calcula la exactitud de la prueba, comparando el nivel de precisión con respecto al patrón de oro de escala ordinal. En el estudio de la variable IRM se tuvo en cuenta el tipo de lesión que se observa en el paciente: clásica y no clásica. Las definiciones son:

- Clásicas (las que ya se consideran expresión de Enfermedad Cerebrovascular Asintomática ( Leucoaraiosis, hiperintensidades, infartos lacunares).
- 
- No clásicas ( se proponen como marcadores de daño vascular aunque su aparición puede tener otras causas (Dilatación de los espacios de Virchow Robbins y la Atrofia) (Johann 2009).

Los hallazgos de lesiones en las imágenes de resonancias se expresan a través de las variables IRM en la figura 1.



**Figura 1.** Resultados de los estudios de IRM de cráneo

Durante el cálculo del área bajo la curva ROC, se utilizaron varias matrices de penalización. El estudio realizado llevó a determinar que la de mejor comportamiento fue:

$$M = \begin{pmatrix} 0 & 0,5 & 1 \\ 0 & 0 & 0,5 \\ 0 & 0 & 0 \end{pmatrix}$$



Esta asigna una menor penalización a los estados de verdad más cercanos.

La tabla siguiente muestra las variables que alcanzaron mayor valor de área bajo la curva, lo que implica que tienen un mayor el nivel de precisión para la inferencia de presencia o no de lesiones cerebrales. Note que los errores estándar son muy pequeños.

Realizando un análisis similar se considera a continuación la comparación del modelo ÁUC (Estándar no Binario contra el Binario). En este caso el estadístico  $\theta_1$  representa el no binario y  $\theta_2$  binario en una prueba. Nuestra hipótesis ahora es  $H_0: E(Z) \leq 0$  vs  $E(Z) > 0$

VARIABLES NEUROFISIOLÓGICAS	AUC(ROC)=ORD(IRM, VARIABLES NEUROFISIOLÓGICAS)	ERROR STANDARD
Grand Total EEG (GTE) score	0.7894737	0.04530824
Coficiente de inteligencia ejecutivo	0.7695838	0.04535125
Organización de la Percepción	0.7619339	0.04376077
Figuras incompletas	0.8115055	0.03821995
Bloques	0.7695838	0.04645697
Aritmética	0.7628519	0.04552472
Ordenamiento de dibujos	0.7925337	0.0446074
Sucesión de letras y números	0.7616279	0.04770905
Ensamble de objetos	0.7937576	0.04279166

**Tabla 3.** Comparaciones entre las AUC en las cuales los contrastes no difirieron significativamente.

Nuevamente usamos el estadístico de (Hanley 1983) presentado anteriormente y se muestran los resultados en la tabla 3.

PRUEBAS	ACB Estándar No binario	ÁCB Estándar Binario	RESULTADO DE LA COMP. DE LOS MODELOS	PRUEBAS	ACB Estándar No binario	ÁCB Estándar binario	RESULTADO DE LA COMP. DE LOS MODELOS
Fondo de ojo	0.7433	0.5768	$z = 3.0202$ $p = 0.0042$	Coficiente de inteligencia verbal	0.71695	0.5526	$z = 2.9810$ $p = 0.0047$
Minimental	0.71909	0.5767	$z = 2.5818$ $p = 0.0142$	Coficiente de inteligencia ejecutivo	0.76958	0.6318	$z = 2.5000$ $p = 0.0175$
Dopp.Carot	0.73623	0.5608	$z = 3.1819$ $p = 0.0025$	Comprensión Verbal	0.72644	0.5768	$z = 2.7150$ $p = 0.0100$
Gran Total EEG (GTE) score	0.78947	0.6824	$z = 1.9421$ $p = 0.0605$	Organización de la Percepción	0.76193	0.5789	$z = 3.3191$ $p = 0.0016$

**Tabla 4.** Representación de los resultados de las comparaciones de la precisión de algunas variables que muestran el funcionamiento del cerebro mediante los dos modelos (Estándar no Binario contra el Binario) .

Las pruebas hacen una partición de los pacientes Vale preguntarnos si hay una similitud entre dos biomarcadores. Hay una extensa literatura sobre índices de similitud.

La partición de los pacientes en afectados y no afectados hecha por un biomarcador, al compararse con el resultado de RM es un segundo criterio de evaluación de estos. Al analizar las particiones tenemos una representación en Tablas de contingencia, tabla 5.

Note que si  $n$  es el número de pacientes analizados  $N$  es de pares de resultados ( $RM, BM$ ).

Albetineh et al (2006) definieron la clase de índices de similitud

$$\mathcal{L} = \left\{ \text{Indice} = \alpha + \beta \sum_{i=1}^I \sum_{j=1}^J m_{ij}^2, \alpha, \beta \text{ \u00fanicos} \right\}$$

Las variables  $m_{ij}$  son las entradas en una tabla  $I \times J$ . En esta familia tenemos los siguientes \u00edndices

$$I_R = \frac{A+B}{N}, \alpha = 1 - \frac{\sum_{i=1}^I m_{i+}^2 + \sum_{j=1}^J m_{+j}^2}{n(n-1)}, \beta = \frac{2 \sum_{i=1}^I \sum_{j=1}^J m_{ij}^2}{n(n-1)}$$

$$I_{CZ} = \frac{2A}{2A+B+C}, \alpha = \frac{-2n}{\sum_{i=1}^I m_{i+}^2 + \sum_{j=1}^J m_{+j}^2 - n}, \beta = \frac{\sum_{i=1}^I \sum_{j=1}^J m_{ij}^2}{\sum_{i=1}^I m_{i+}^2 + \sum_{j=1}^J m_{+j}^2 - n}$$

		BIO-MARCADOR (BM)		
RESONANCIA MAGN\u00c9TICA (RM)	GRUPO DE CLASIFICACI\u00d3N	AFECTADO	NO AFECTADO	TOTAL
	AFECTADO	$A(m_{11})$	$B(m_{12})$	$A+B(m_{1+})$
	NO AFECTADO	$C(m_{21})$	$D(m_{22})$	$C+D(m_{2+})$
		$A+C(m_{+1})$	$B+D(m_{+2})$	$N=C^n_2$

**Tabla 5.** Resultados de BM vs RM

Analizaremos la similitud existente entre los biomarcadores con la RM. Utilizamos estos dos \u00edndices y se obtuvieron los resultados en la tabla 6.

Como se ve sigue siendo las figuras incompletas el de mejor para identificar los resultados de RM. Le sigue en esta capacidad GTE que ten\u00eda un comportamiento no distintivo entre las variables seleccionadas. El ensamble y el ordenamiento de dibujos mantienen su nivel de importancia.

VARIABLES NEUROFISIOL\u00d3GICAS	$I_R$	$I_{CZ}$
Gran Total EEG (GTE) score	0,9205	0,9524
Coficiente de inteligencia ejecutivo	0,9004	0,8724
Organizaci\u00f3n de la Percepci\u00f3n	0,8670	0,8218
Figuras incompletas	0,9763	0,9803
Bloques	0,8400	0,8174
Aritm\u00e9tica	0,7994	0,8217
Ordenamiento de dibujos	0,9129	0,9342
Sucesi\u00f3n de letras y n\u00fameros	0,8763	0,8185
Ensamble de objetos	0,9573	0,9523

**Tabla 6.** Similitud existente entre los biomarcadores con la RM

Una comparaci\u00f3n entre el uso del est\u00e1ndar binario y el no binario es el dado en la tabla 7.

Como se ve nuevamente el Gran Total es el que exhibe mejor comportamiento. Aun as\u00ed no se puede decir que el uso de ambos est\u00e1ndares genere resultados similares. Esto es consecuente con el hecho de que al comparar la precisi\u00f3n aceptando la aproximaci\u00f3n normal planteada por Obuchowski (2005)

$$F \left( \frac{\hat{\theta}^1 - \hat{\theta}^2}{\sqrt{\text{Var}(\hat{\theta}^1) + \text{Var}(\hat{\theta}^2) - 2\text{Cov}(\hat{\theta}^1, \hat{\theta}^2)}} \middle| \theta^1 = \theta^2 \right) = 0,0605$$

PRUEBAS	$I_R$	$I_{CZ}$	PRUEBAS	$I_R$	$I_{CZ}$
Fondo de ojo	0,6984	0,7073	Coficiente de inteligencia verbal	0,7051	0,7125
Minimental	0,7152	0,7247	Coficiente de inteligencia ejecutivo	0,8727	0,8099
Dopp.Carot	0,6267	0,8845	Comprensión Verbal	0,6658	0,6983
Grand Total EEG (GTE) score	0,8533	0,8416	Organización de la Percepción	0,6349	0,5962

**Tabla 7.** Similitud existente entre el estándar binario y no binario

La varianza es estimada mediante

$$Var(\hat{\theta}_{il}) = \sum_{t \in I} \sum_{s \in I} \sum_{l \in I} \sum_{i \in I} W_{ts} \cdot W_{il} \cdot L_{ts} \cdot L_{il} \cdot cov(\hat{\theta}_{ts}, \hat{\theta}_{il})$$

donde

$$cov(\hat{\theta}_{ts}, \hat{\theta}_{il}) = \begin{cases} Var(\hat{\theta}_{ts}) & \text{si } t = i, s = l \\ \frac{1}{n_s} S_t^{ts, is} & \text{si } t = i \\ 0 & \text{si } t \neq i, l, \text{ y } s \neq i, l, \\ \frac{1}{n_t \cdot n_l - 1} \sum_{j \in I} [L_{jt} \cdot \hat{\theta}_{jt} - L_{jl} \cdot \hat{\theta}_{jl}] \cdot [L_{js} \cdot \hat{\theta}_{js} - L_{jl} \cdot \hat{\theta}_{jl}] & \text{en otro caso} \end{cases}$$

## 5. CONCLUSIONES

- El uso de BM's como sustitutos del uso de RM es recomendable dado su impacto económico y porque sus resultados son compatibles desde el punto de vista del comportamiento de su capacidad para detectar daños cerebrales.
- 
- Del grupo de BM's clasificados como aceptables por La Rosa (2011) tiene una mayor similitud con los resultados de la RM Variables Neurofisiológicas y Figuras incompletas.
- 
- Del análisis de la similitud se desprende que las variables asociadas al uso de habilidades lógico-manuales son las más importantes: Figuras incompletas, Ensamble de objetos, Ordenamiento de dibujos.
- 
- De medirse todas las variables el Gran-total permite caracterizar la similitud con el resultado del uso del RM. Valdría la pena hacer el análisis de que ponderaciones podrían generar un índice que mejorase el GTE actual.
- 
- Sigue siendo válido las conclusiones de La Rosa (2011) de que no se puede utilizar indistintamente el estándar binario y el no binario.

## REFERENCIAS

- [1] ALBATINEH, A.N. & M. NIEWIADOMSKA.BUGAJ (2011): Correcting Jaccard and other similarity indexes for chance agreement in cluster analysis. **Adv. In Data Analysis and Classif.** 5, 179-200.
- [2] ALBON, E., TSOURAPAS A, FREW E, DAVENPORT C, OYEBODE F, BAYLISS S, ARVANITIS T. & MEADS C (2008): Structural neuroimaging in psychosis: a systematic review and economic evaluation. **Health Technol Assess** 12, 163-176.
- [3] BIASUTTI, M., DUFOUR N, FERROUD C, DAB W. & TEMIME L (2012): Cost-Effectiveness of Magnetic Resonance Imaging with a New Contrast Agent for the Early Diagnosis of Alzheimer's Disease. **PLoS ONE** 7(4): e35559. doi:10.1371/journal.pone.0035559

- [4] BOUZA, C. & M. Li. (2006) : The use of the receiver operating characteristic in the retrieval of documents in a web data base. **Revista Investigación Operacional** 27, 300-315
- [5] CARMEN M, S. C. (2010) : **Metodología ROC en la Evaluación de Medidas Antropométricas como Marcadores de la Hipertensión Arterial**. Santiago de Compostela, Universidad de Santiago de Compostela
- [6] CHAP , T. L. (2006) : A solution for the most basic optimization problemas sociated with an ROC curve. **Statistical Methods in Medical Research** 15, 571-584
- [7] HANLEY, J. A. &MCNEIL, B.J. (1983) : A method of comparing the areas under receiver operating characteristic curves derived from the same cases. **Radiology** 148, 839-843.
- [8] LA ROSA T. (2011): **EVALUACIÓN MEDIANTE CURVAS ROC DE LAS POTENCIALIDADES DE MARCADORES PRECOCES DE LESIÓN CEREBRAL EN PACIENTES HIPERTENSOS NEUROLÓGICAMENTE ASINTOMÁTICOS**. Tesis de Maestria , Universidad de La Habana.
- 9] LE, C. T. (2006) : A solution for the most basic optimization problem associated with an ROC curve. **Stat.Methods Med.Res.** 15, 571-584.
- [10] LUSTED, L. B. (1971) : Decision making studies in patient management. **N Engl J Med** 284, 32-38.
- [11] OBUCHOWSKI, N. P. (2005) : Estimating and Comparing Diagnostic Tests' Accuracy When the Gold Standard Is Not Binary1. **Academic Radiology** 12, 67-78.
- [12] PEPE, M. S. (1998) : Three approaches to regres analysis of receiver operating characteristic curves for continuous test results. . **Biometrics** 54, 124-135.
- [13] STADLBAUER, A ; BERNT R. ; SALOMONOWITZ E. ; PLAS E. ; STRUNK G. &EBERHARDT K. (2011): Health-Economic Evaluation of Magnetic Resonance Imaging Before Biopsy for Diagnosis of Prostate Cancer. **RöFo. Fortschritte auf dem Gebiet der Röntgenstrahlen und der bildgebenden Verfahren** 183, 925-932
- [14] ZHOU, X. C., P. (2003) : Nonparametric Analysis for the ROC Area of two Diagnostic Test in the Presence of Nonignorable Verification Bias. **Statistics Planning and Inference** 115, 193-213.
- [15] ZWEIG, M. H. & CAMPBELL, G. (1993) : Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. **Clin Chem** 39, .342-356